FROM NARROWBAND TO BROADBAND:
ASYNCHRONOUS TRANSFER MODE
AND THE
TRANSFORMATION OF THE TELEPHONE NETWORK

by

STEPHEN JOHANNES DOWNS

B.S., Physics and Applied Physics, Yale University (1986)

Submitted to the Department of Electrical Engineering and Computer Science
and the Technology and Policy Program
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN TECHNOLOGY AND POLICY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August, 1992

Signature of Author ______________________________
Department of Electrical Engineering and Computer Science
August 10, 1992

Certified by ______________________________
Dr. Lee W. McKnight
Research Associate, Center for Technology, Policy, and Industrial Development
Thesis Supervisor

Accepted by ______________________________
Professor Richard de Neufville
Chairman, Technology and Policy Program

FROM NARROWBAND TO BROADBAND:
ASYNCHRONOUS TRANSFER MODE
AND THE
TRANSFORMATION OF THE TELEPHONE NETWORK

by

STEPHEN JOHANNES DOWNS

Submitted to the Department of Electrical Engineering and Computer Science
and the Technology and Policy Program
on August 10, 1992
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Technology and Policy

ABSTRACT

The public telephone network has evolved, in its capabilities and services, to the point where it is no longer merely a telephone network. The capability of optical fiber transmission to support high-speed, or large bandwidth, communications has led to the conception of *broadband* communications networks, as opposed to the low-speed, narrowband facilities offered by the traditional public telephone network. This progress, from narrowband to broadband networks, promises to transform fundamentally the nature of the public telephone network and thereby threatens to make obsolete many of the policies that have served the telephone industry since the birth of the Bell System.

The development of integrated broadband networks requires an entirely new approach to network architecture and switching principles. The approach chosen by the world's telecommunications carriers is known as Asynchronous Transfer Mode, or ATM. This thesis examines the principles and specifications of ATM and probes the nature of ATM-based public broadband networks.

ATM-based networks can be regarded as service-independent communications platforms. With the introduction of ATM technology, the telephone network will be transformed to the point that telephone service will become just one of many communication services provided on top of a more fundamental network. Because telephone service and basic network provision will have different economic characteristics and because they may not share the same policy goals, telecommunications policy should reflect this conceptual separation of the telephone network from the more basic platform network.

An examination of the structure of ATM-based broadband communications networks shows that their functions are both separable and modular. This functional architecture offers an opportunity for the furtherance of the Federal Communications Commission's policy of Open Network Architecture. The degree to which the network will ultimately be open will depend strongly on the results of ongoing international standardization efforts.

Thesis Supervisor: Dr. Lee W. McKnight

Title: Research Associate,
Center for Technology, Policy, and Industrial Development

ACKNOWLEDGMENTS

I would like to thank Dr. Lee McKnight, my thesis supervisor, for all of his guidance, suggestions, and encouragement. I am particularly grateful to Dr. McKnight for giving me the opportunity to explore a topic of my own choosing. I would also like to thank Prof. David Tennenhouse, of the MIT Laboratory for Computer Science, for reviewing early drafts of the thesis and providing valuable, insightful comments.

I wish to thank the following people for providing helpful comments along the way, pointing me in the direction of useful sources, and assisting me in tracking down various items: Jane Bortnick, David Reed, Evan Kwerel, Gerry Waldron, Colin Crowell, Robert Smith, Manley Irwin, Francis Dummer Fisher, Susan Baldwin, Audrey Bashkin, Richard Solomon, Suzanne Neil, Jan van Ruymbeke, Wim De Meyer, and Sara Verhas.

My introduction to Asynchronous Transfer Mode technology came at the Alcatel Bell Research Center in Antwerp, Belgium as part of an internship with the Belgian RTT. I am deeply grateful to Dr. Martin De Prycker for hosting my visit to Alcatel Bell and arranging for my participation in a seminar on ATM technology. I am also grateful to my hosts at the Belgian RTT, especially Maxime Watteyne, for exposing me to the Belgian Broadband Experiment, the RACE program, and other European broadband initiatives. Thanks must also go to Susan Mirbach of Belgian Telecom USA and the ICA Foundation, which sponsored my internship.

Special thanks must go to Janet Estes, for her continual encouragement, and to my parents, for all of their support. I am especially grateful to my father for his always constructive criticism and for his genuine interest in my research. Finally, I wish to thank my colleagues and friends in the Technology and Policy Program, especially the EFS gang, and the staff of the Center for Technology, Policy, and Industrial Development for providing such an enjoyable research atmosphere.

This research has been sponsored in part by a grant from the Defense Advanced Research Projects Agency, contract number N000174-92-C-0020.

Table of Contents

List of Figures

Chapter One

Introduction

1.1 INTRODUCTION

The public telephone network has evolved, in its capabilities and services, to the point where it is no longer merely a telephone network. The adoption of digital switching and transmission techniques and the growing use of optical fiber as a transmission medium enable the communication of virtually all forms of information over the network. The capability of optical fiber transmission to support high-speed, or large bandwidth, communications has led to the conception of *broadband* communications networks, as opposed to the low-speed, narrowband facilities offered by the traditional public telephone network. This progress, from narrowband to broadband networks, promises to transform fundamentally the nature of the public telephone network and thereby threatens to make obsolete many of the policies that have served the telephone industry since the birth of the Bell System.

Providing a multitude of different services over an international network of interconnected networks requires coordination among the telecommunications providers of the world. This coordination traditionally comes from the International Telegraph and Telephone Consultative Committee (CCITT) of the International Telecommunication Union, a United Nations-sponsored treaty organization (Codding and Rutkowski, 1982). In recognition of the potential of evolving telephone networks and the diverse demands of communications users, the CCITT has arrived at a series of recommendations on the architectural and technological principles around which future broadband public networks will be organized. The CCITT has a vision of a single, unified network that supports a very wide range of communication services over a limited set of interfaces—a "one size fits all" approach to communications. The realization of this ambitious vision for broadband networks requires a new approach to network architecture and switching principles. The broadband networks of the future will be achieved, not by the extension of the current network, but by the construction of an entirely new foundation.

The CCITT's vision is known generally as ISDN, for Integrated Services Digital Network.[1] The application of ISDN principles to broadband, optical fiber networks is termed *Broadband* ISDN, or B-ISDN. In 1988, the CCITT selected a technique known as Asynchronous Transfer Mode, or ATM, as the target solution for Broadband ISDN.[2] While previous enhancements to the telephone network, including early implementations of ISDN, can be described as evolutionary, a restructuring of the network around ATM principles is truly revolutionary. ATM represents a total change in the philosophy of how network communications capacity is to be allocated. Communication channels, instead of being rigidly defined to support standard voice calls, will be flexibly assigned to accommodate the individual characteristics of each communications application. The choice of ATM is also highly significant in that it represents an attempt to generalize the functionality of the network. The public telephone network has always been optimized to suit the particular characteristics of telephone communication. By selecting ATM as the solution for broadband networks, the CCITT chose not to optimize the performance of the network around the needs of any one particular communication service. Instead, it opted for flexibility, choosing a solution that could accommodate virtually all communication services.

Broadband ISDN and the adoption of ATM switching principles into the public network must not be seen as an enhancement of the current network. While ATM will enable an

---

[1] For an overview of the principles and history of ISDN, see Rutkowski (1985).

[2] Minzer (1989) provides a general sketch of B-ISDN and ATM principles.

expansion of the service offerings of the carriers, this expansion will result, not from a grafting of new capabilities onto current capabilities, but from a fundamental restructuring of the network itself. Although it is far from certain that a public switched broadband network, analogous to the public switched telephone network, will become a reality, an exploration of the nature of such a network is warranted. The prospects for a widely available public network based on ATM will be influenced by both market forces and public policy. Public policy decisions concerning the promotion or influence of network development, must, however, be informed by an understanding of the qualities of any future public network. The nature of a new, ATM-based public network—its capabilities, its organization, and its implications for policy—is therefore the subject of this thesis.

1.2 STATEMENT OF THE THESIS

The transformation of the public network, from a telephone network to a general, all-purpose communications network, will both necessitate new policy approaches and create opportunities for the furtherance of existing policy goals. In the past, the telephone network evolved gradually to accommodate new communications applications. In contrast, the organization of the new network will be such that telephony will not be part of the basic network service—it will be one of many possible applications of a more fundamental network service. As the provision of telephone service and basic network service will be subject to different technical and economic characteristics, they will require separate policy approaches. The restructuring of the network will also present policy makers with opportunities to further the goals of network openness and innovation in enhanced services. The functions of the network, it will be shown, are being organized into modular, separable components. This approach to network design could, if policy goals are incorporated into the international standards process, result in a truly open architecture for public broadband networks.

1.3 OVERVIEW OF THE THESIS

The thesis is organized into three main parts, covering the evolution of the public network and the policy response to its growing capabilities, ATM and its potential effects on the structure of the public network, and policy opportunities presented by the transformation of the network.

Chapter 2 describes the technological and service evolution of the public network, showing how it has progressed from a single service network to one that can support multiple communication services. Innovations in internal network technology are highlighted along with innovative uses of the public network. Chapter 2 also contains a discussion of the response of the telephone companies to the growing demand for diversified, non-telephone services.

A historical sketch of the policy responses to the evolution of the network is offered in Chapter 3. The chapter is divided into two parts, focusing first on the establishment of a single, unified, public telephone network and later on the regulatory responses to the advent of computer communications and the blurring of distinctions between telephony and other applications.

Chapter 4 is an overview of the CCITT's approach to Broadband ISDN. A discussion of the services envisioned for B-ISDN is followed by an analysis of the challenges inherent in developing a single network capable of supporting those services. The technical principles, features, and capabilities of the solution, Asynchronous Transfer Mode, are subsequently reviewed.

Chapter 5 is the first of two chapters analyzing some of the fundamental characteristics of public ATM networks. The focus of Chapter 5 is on the service-independent nature of an ATM-based network. Application scenarios, demonstrating how telephony and other applications could be offered upon an ATM platform, are presented.

Chapter 6 consists of an enumeration and description of the various functions provided by an intelligent broadband network and an organization of these functions into a conceptual model of the network. The separability and modularity of these functions is highlighted.

The prospects for a widely available public ATM network are assessed in Chapter 7. Public ATM networks are compared and contrasted with the current public switched telephone network and ATM is discussed with reference to alternative broadband networks, such as cable television distribution systems.

Chapter 8 concludes the thesis by offering three general policy recommendations. First, it is argued that policy makers should make a distinction between telephone service and the underlying basic network services that will form the foundation for future telephone networks. Second, the opportunities for the design of future public broadband networks along open architecture principles are presented. It is recommended that the Federal Communications Commission participate in the ongoing Broadband ISDN standardization efforts for the purpose of furthering its Open Network Architecture policies. Finally, possible approaches to universal service are discussed.

Chapter Two

The Technological Evolution of the Network

2.1 INTRODUCTION

Over the past thirty years, the public telephone network[3] has evolved to the point where it is capable of providing communication services beyond what the industry terms "Plain Old Telephone Service," or "POTS." A combination of technological breakthroughs and creative uses of the network have fueled this evolution. Innovations in network technology, such as the advent of digital transmission techniques and the development of optical fiber as a communications medium, that now enable the provision of diverse communications services were not developed for the purpose of offering new services, but were seen as ways to increase the internal efficiency of the telephone network (Bolter and McConnaughey, 1991). Parallel to these internal developments, the user community and equipment providers developed non-telephony communications applications that could utilize the public telephone network, despite its orientation and optimization for the transmission of voice signals. The telephone companies have responded to the growing demand for non-telephony communications with plans both to integrate data communications within the telephone network and to offer communication services using separate, overlay networks.

---

[3] The term "public telephone network" is used in this context, rather than the more conventional "public switched telephone network," to include all of the facilities, including unswitched transmission links, of the public switched network.

2.2 INTERNAL DEVELOPMENTS

Alexander Graham Bell invented the telephone using analog principles, suited to the transmission of the human voice.[4] Voices were carried along telephone wires as electrical signals whose amplitudes corresponded to the sound waves produced by the speaker. The telephone network was developed around this method of relaying voices—it has been optimized for the switching and transmission of analog signals comprising a narrow range of frequencies. As a result, the network, when applied to other forms of communication, such as the exchange of computer data or the transmission of images, presents a number of limitations. First, it is designed to allow the passage of voice frequencies (300 Hz to 4 kHz) from end to end, but no others. Any other form of information to be sent across the network must be coded within this frequency band.

The second limitation presented by the telephone network is the method by which calls are switched. Because, traditionally, the only service provided by the network was telephony, the network only offers one type of a channel—a voice channel. Any call, therefore, is established by the temporary provision of a fixed bandwidth voice channel between two endpoints. As the two users converse, this channel, or *circuit*, is dedicated to their conversation--it is generally not shared with other users. Any excess bandwidth in the channel, made available by pauses in the conversation, is essentially wasted by the network. This type of switching, in which a full, fixed bandwidth channel is dedicated to each call, is known as circuit switching. Circuit switching, while generally appropriate for voice communications, is highly inefficient for communications where the information transmitted is generated in bursts, as is typical for many computer communications.

The telephone network has evolved considerably since World War II, capitalizing on the invention of the transistor and the ensuing electronic revolution. The application of new

---

[4] The use of analog signals was actually a departure from the techniques employed by the telegraph, which involved the transmission of digitally coded text messages.

technologies to the network have for the most part been undertaken primarily for the purpose of improving the efficiency of providing telephone service, as opposed to the expansion of the network's service capabilities. However, two significant changes in the network, brought about by efficiency considerations, will enable a shift in the orientation of the telephone network into an all-purpose communications network.

2.2.1 The Digitalization of Transmission

The digitalization of the telephone transmission network began in 1962, with the first use of a digitally encoded T1 carrier system in the public network (O'Neill, 1985).[5] Voice signals were converted to streams of digital data using a technique called Pulse Code Modulation (PCM). PCM, in which an analog voice signal is sampled periodically and discrete values assigned to its amplitude, had been conceived many years before, but never implemented in the telephone network due to cost considerations. Indeed, for many years at Bell Laboratories, PCM had been considered a solution looking for a problem. The problem that eventually arose to meet this solution was the growth in inter-office traffic within the local exchange network. As local operating companies were forced to install more and more cable to support calls that traveled between central office switches, they sought cost-effective solutions for multiplexing calls over existing facilities. By the late 1950s, advances in solid-state electronics had driven the cost of PCM coding to the point where it could be considered a practical alternative to analog solutions.

The introduction of PCM-based, digital transmission into the exchange network proved successful and digital techniques soon spread rapidly from the exchange networks to the long-distance transmission networks. Distribution cable in the subscriber loop network is now frequently augmented with digital PCM multiplexing equipment (Shumate and Snelling, 1991). While the purpose of introducing digital transmission into the network was to achieve a more

---

[5] A T1 carrier system uses time-division multiplexing to carry twenty-four 64 Kilobit per second (Kbps) channels on copper wire or coaxial cable.

efficient utilization of the network's transmission plant, the creation of a digital transmission infrastructure offered opportunities for the carriage of digital information besides PCM voice. When this digital infrastructure was extended all the way to the user in the early 1980s, the telephone network took on a new role. High-speed transmission channels, available for private applications, were offered by the telephone company unbundled from any telephone service. The low-level transmission infrastructure of the telephone company, insofar as it could be made available for private, non-telephony uses, could be separated, conceptually, from the telephone network itself.

2.2.2 The Introduction of Optical Fiber

Bell Laboratories' invention of the laser in 1960 sparked numerous research efforts in the field of optical communications (Li, 1983). Ten years later, Corning Glass Works manufactured an optical fiber with signal loss characteristics sufficiently low to make fiber transmission of telephone signals economically feasible. Transmission over optical fiber offered the advantages of using lightweight cables, being immune to electromagnetic interference, supporting enormous communications capacity, and needing fewer signal regenerations than would transmissions over copper or coaxial cables. These characteristics made optical fiber particularly suitable for long-distance transmissions. Once low loss transmission had been demonstrated in the laboratory, development work on optical fiber systems began, and by 1979, a 45 Megabits per second (Mbps) DS-3 carrier, bundling 28 T1 carriers for a total of 672 voice circuits, was installed commercially (O'Neill, 1985).

The performance of optical fiber transmission continues to improve. Signal loss is below a single decibel per kilometer (Corning's 1970 breakthrough was a 20 dB/km fiber) and higher and higher data rates continue to be realized. Optical transmission systems that operate at 2.5 Gigabits per second (Gbps) are now widely available (Kaiser, 1991). Transmission at a rate of 20 Gbps for 10,000 km without repeaters has been demonstrated in laboratory experiments (Lucky, 1992). Optical fiber transmission is now the norm for both local exchange and long-distance traffic and is beginning to appear in the subscriber loop network (Shumate, 1989). Transmission over optical fiber, like digital transmission, was pursued as an opportunity to reduce the cost of transmitting voice channels. However, as the fiber moves closer to the user, it will offer interfaces of extremely high data rates, orders of magnitude greater than is provided for voice channels. These high-speed interfaces will enable the provision of bandwidth-intensive communication services, namely those involving high-resolution, full-motion video, that cannot be realized over traditional copper loops.

2.3 ALTERNATIVE APPLICATIONS OF THE PUBLIC NETWORK

While AT&T and the independent telephone companies were improving the efficiency of the network, users and equipment manufacturers began to develop new applications for the network. The advent of digital computers led to a demand for computer communications and out of this demand came the invention of the modem. As the nature of computer communications became better understood, independent networks, optimized for the exchange of computer data, were built on top of the public network. With the popularization of the fax machine, the 1980s saw an explosion in the use of the network for image communications.

2.3.1 The Modem

A modem, which derives its name from its dual functions of modulation and demodulation, converts digital computer data to analog signals at one end of a connection and converts analog signals to digital data at the other. In order for communications to pass through the public telephone network, a modem must use voice band frequencies. This requirement, to stay below 4 kHz, imposes a practical limitation on the speed at which data can be transmitted over analog telephone lines.

While techniques for transmitting digital data over analog facilities had been in existence since the early days of telegraph communication, serious development of modem technology began in the 1950s as part of the Department of Defense's SAGE (Semiautomatic Ground Environment) project (Radford, 1962; O'Neill, 1985). SAGE, a complex national air defense system, required the transmission of data from remote radar stations to central computer processing units. Research and development efforts at both Bell Laboratories and MIT Lincoln Laboratory led to the development of modems that could carry data, over specially tailored private lines, at the relatively high data rate of 1200 bits per second (O'Neill, 1985). Commercial modems supporting communications over switched, public facilities were subsequently offered by AT&T in 1960. Today, over 2.5 million modems are sold in the United States each year (Electronic Industries Association, 1992).

The modem, by converting data into voice-band signals, has had the effect of stretching the capabilities of the public telephone network beyond the limits of telephone service. While the telephone network has limited communications capabilities—only voice-band signals can pass through—the modem allows users to treat the network as an all-purpose communications network. As regulatory policies (to be described in Chapter 3) gradually removed AT&T's control over terminal equipment, the telephone network ceased to be a telephone network so much as a communications network that was optimized to provide telephone service. The actual type of information to be communicated, be it voice or data, was left to the user.

2.3.2 Packet-Switched Networks

Modems made computer communications over the telephone network possible but far from ideal. Computer communications, quite simply, have different characteristics than voice communications. First, while voice communications are rather steady (depending, of course, on the speakers), computer communications are often "bursty," coming in spurts. Fixed rate channels are not optimally suited for bursty transmissions because they often do not provide enough bandwidth when it is needed and because, during the "silences" between computer transmissions, the bandwidth of the channel is wasted. Furthermore, while for voice conversations to be intelligible, transmission delays must be consistent (and small), most computer communications can tolerate variable delays.

In the 1960s, as AT&T was investing heavily in the development of its new electronic telephone circuit switches, it had little incentive to develop a new approach to switching that would better suit data communications. That push came instead from the Department of Defense, which, as part of a project to connect computers located at research universities, implemented the first network to use the principle of packet switching (Roberts, 1978; Cerf, 1991). Packet switching, as distinguished from circuit switching, allocates channel bandwidth dynamically. Data, as it is generated, is assembled into discrete packets that are routed through the network individually, according to their address labels. Packets associated with multiple calls are intermingled, traveling through shared circuits between packet switches. At each switching node, a packet's label is read and processed, and the packet is routed to the next node, along the path to its intended destination. Efficiencies are gained by capitalizing on the statistical nature of most data transmissions. In many applications, packets are not generated continuously, but sporadically, in "bursts." Network facilities, such as the fixed rate, unswitched channels that connect packet switches, are dimensioned to carry the aggregate traffic of a number of bursty sources rather than the sum of the peak data rates required by each source. Packet-switched networks also hold an advantage in that they can provide a graceful degradation of service. If all of the facilities of a circuit-switched network are in use, calls cannot get through—a busy signal is returned and the call request denied. In a packet-switched network, congestion leads to slower performance—packets are delayed in the network—but in most cases, they do get through.

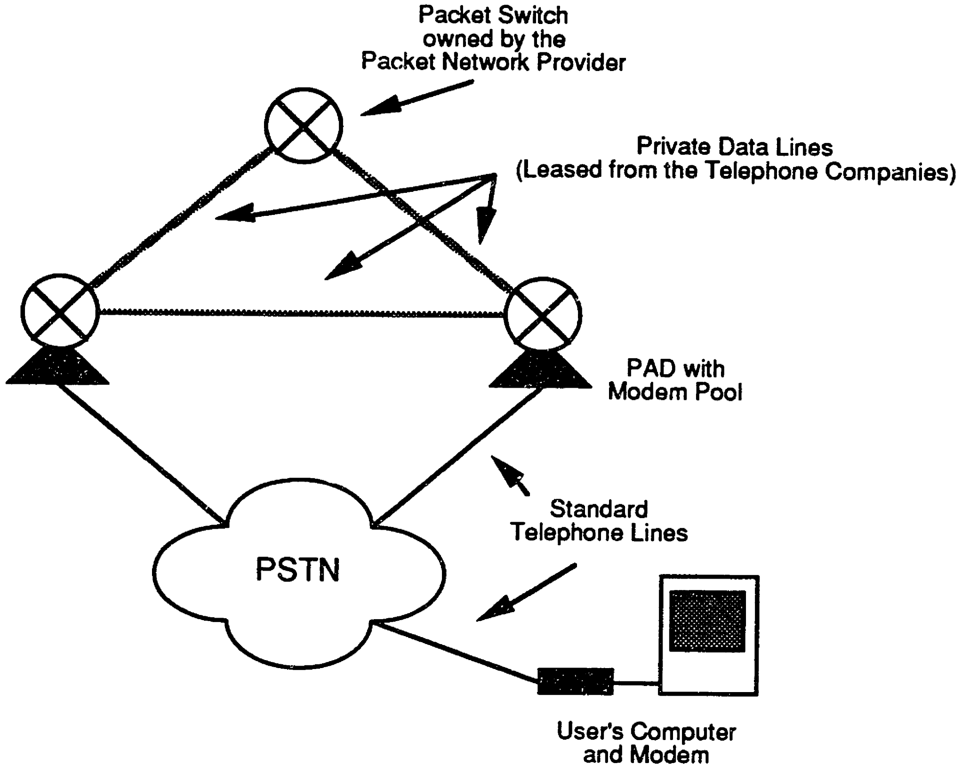
The Defense Department's packet network, known as the ARPANET after the responsible organization, the Advanced Research Projects Agency, was inaugurated in 1969 with four nodes and grew rapidly into what is now the multinational Internet,[6] comprising thousands of network nodes. In the early 1970s, ARPA approached AT&T to determine whether they would be interested in taking over the ARPANET, operating it as a public packet-switched service, but AT&T declined. The telephone company's decision opened the door for the formation of a number of public packet carriers, most notably Telenet, a spin-off of the company that held the original ARPANET contract, and Tymnet, a computer time-sharing company. These public packet-switched networks were provided by utilizing selected portions of the public telephone network (see Figure 2.1). The packet carriers' packet switches were interconnected using leased, unswitched data circuits from the telephone company. User access to the packet carriers often came through dial-up connections made over the public switched telephone network (PSTN). A user with a modem would dial a local or regional telephone line owned by the packet carrier. This access point connected the user with a packet assembler/disassembler (PAD) that converted the asynchronous data transmissions of the user into packets that could be routed across the network. Through access to the packet network, one could connect to other users, to computer bulletin board services, and to other services, often provided by the packet network provider itself, such as electronic mail.

The advent of packet switched networks established the value of the public telephone network as a form of infrastructure. Different components of the network, dedicated transmission lines and standard switched services, could be combined to provide what amounted to higher level networks. These higher level network services could be offered directly to customers and were beyond the control of the telephone companies providing the underlying network capabilities.

---

[6] See Chapter 7 for a description of the Internet.

Figure 2.1.
Simplified architecture of a public packet-switched data network.



2.3.3 Image Communications

Another non-telephony use of the PSTN that has become increasingly common in the last ten years is the transmission of image facsimiles. The fax machine, little more than a modem combined with an image scanner, has quickly pervaded our national culture.[7] Facsimile images transmitted over the PSTN are in black and white and limited to relatively low resolution whereas networks with greater transmission capabilities could support color and higher resolution images. The rapid acceptance of PSTN-based fax, despite the network's shortcomings,

---

[7] The North American Telecommunications Association found that in 1989, fax machines were in use in 97% of all manufacturing facilities and in 96% of all business service organizations (Electronic Industries Association, 1991).

suggests the power of a ubiquitous communications network. For fax, as for computer communications, the PSTN serves as a platform for applications development, providing limited capacity, but offering a standard interface and connectivity to virtually the entire population.

2.4 TELEPHONE COMPANY RESPONSES TO DEMAND FOR DIVERSIFIED SERVICES

The telephone companies' primary response to the proliferation of non-telephony communications applications has been the Integrated Services Digital Network, or ISDN. ISDN represents a vision of a ubiquitous digital communications network, providing standard interfaces, across which all forms of information could be transmitted. The term "ISDN" was coined in the early 1970s by the CCITT, but serious development work to flesh out the vision did not begin until 1980 (Rutkowski, 1985). Recommendations for ISDN principles and standards were released by the CCITT in 1984 and 1988; ISDN implementation is ongoing.

2.4.1 Narrowband ISDN

The first implementation of the ISDN vision is termed "Narrowband ISDN," or "N-ISDN" to distinguish it from the Broadband ISDN (B-ISDN) currently under development. Rather than the development of a new network, N-ISDN is an attempt to enhance the capabilities of the existing telephone network. N-ISDN uses the same copper wires for subscriber loops and the same central office switches, albeit somewhat augmented, as the PSTN. Whereas the introduction of PCM-based transmission systems brought digital technology into the inter-office and long-distance transmission networks, N-ISDN promises to bring the digital interface directly to the subscriber. The CCITT recommendations for N-ISDN specify two user interfaces: the Basic Rate Interface (BRI) and the Primary Rate Interface (PRI). The BRI offers two 64 Kilobit per second (Kbps) channels, known as "B" channels, upon which any information, including PCM voice, can be transmitted, and a 16 Kbps "D" channel, to be used for control signaling. The PRI, designed primarily for connections to private branch exchanges, specifies twenty-three 64 Kbps B channels and one 64 Kbps D channel.

N-ISDN's digital interfaces connect directly to digital computer data terminals and thus obviate the need for modems. However, despite the digital interface, N-ISDN has been criticized for offering little to the computer user that is not available through the conventional PSTN (Cerf, 1991). The maximum data rate on any channel is 64 Kbps, which, when compared with the 19.2 Kbps transmission available with some modems, is not enough of an improvement for some users to justify investments in new data terminal equipment. Furthermore, N-ISDN, although offering packet switched services, is primarily based on circuit switching, and thus not well-suited to most data communications applications.

Acceptance of N-ISDN has not been widespread. Possible reasons for the lack of enthusiasm for N-ISDN include the inherent inadequacies noted above, a lack of coordination, a problem of availability, and a dearth of compelling applications. Despite years of standardization efforts, implementations of ISDN have featured compatibility problems in the interconnection of local exchange carriers to interexchange carriers and in the interoperability of switches and terminal equipment of different manufacturers. As of June, 1992, one could not make an ISDN data call from one end of the country to another.[8] This lack of interconnection capability is matched by a general lack of availability of ISDN lines. Currently, ISDN is available on only approximately 20% of subscriber telephone lines in the U.S.[9] These problems of incompatibility and availability are currently being addressed by a joint industry effort, spearheaded by Bellcore, the research organization owned jointly by the seven regional Bell operating companies, known as National ISDN-1 (Bellcore, 1991). National ISDN-1 is intended to provide manufacturers and service providers with common technical specifications sufficient for interoperability.

---

[8] "Users Hash Out ISDN Problems," *Communications Week*, June 8, 1992, p. 4, col. 1.

[9] Ibid.

Finally, industry analysts note the lack of what is called a "killer application" for N-ISDN (Baldwin, 1991). A killer application can be understood as one that, by itself, provides sufficient justification or incentive to purchase the underlying product or service. ISDN proponents continue to await an application that could stimulate sales of ISDN service as, for example, the spreadsheet spurred sales of personal computers in the 1980s.

2.4.2 Overlay Data Networks

With the prospects for N-ISDN still uncertain, telephone companies have also sought to meet their customers' data communications needs with high-speed overlay networks. The most highly touted of these non-integrated solutions are known as Frame Relay and Switched Multimegabit Data Service, or SMDS. Both Frame Relay and SMDS are packet-oriented services and require dedicated access lines, separate from telephone subscriber lines. Access speeds for SMDS are projected to fall between 1.2 Mbps and 34 Mbps. Frame Relay interface speeds range from 56 Kbps to 1.5 Mbps.

Frame Relay and SMDS services are supported by standalone broadband switches—they are not integrated into the carrier's telephone central office switches. An analogy can be drawn to the public packet networks discussed earlier in this chapter. Telephone company facilities, such as dedicated transmission lines, are used to interconnect specialized switches, thereby creating the same sort of overlay network. The move toward overlay networks, built upon a common platform of public facilities, marks a recognition among the telephone companies of the growing divergence of user needs. The Broadband ISDN approach discussed below is a long-range attempt to develop an integrated common platform that is flexible enough to meet these ever diverging needs.

2.5 SUMMARY

The public network has evolved internally, through the adoption of digital transmission techniques and the utilization of optical fiber, to the point where it can support communication services well beyond simple telephony. As the demand for other forms of communication has grown, innovators have responded by finding novel ways of utilizing the telephone network to communicate. Their efforts have transformed the telephone network into an all-purpose communications network that is only optimized for telephone service, as opposed to one that is limited to telephone service. Recognizing the demand for non-telephony communications services, the telephone companies have responded with attempts to augment the capabilities of the network. Their first approach was evolutionary, squeezing out extra capacity from the PSTN. The second attempt has been to build new networks on top of the old network. As will be shown in later chapters, the next attempt, Broadband ISDN, will be to build an entirely new foundation.

Chapter Three

The Policy Response

3.1 INTRODUCTION

Chapter 2 presented a picture of the public telephone network as having grown beyond its original function of providing telephone service. The policy response to this development has been complicated by the traditional provision and regulation of telephone service on a monopoly basis. Due to their monopoly status, their conduct, and, in particular, the rates charged by telephone companies have been subject to the review of state and federal regulatory commissions. The rationale for this regulation has been the need for a check on the power of the monopoly to earn unreasonable profits, at the expense of captive ratepayers, without the threat of competition. The regulator is challenged when the monopoly provider ventures forth into potentially competitive markets. While the regulator bears some responsibility for the conduct of the monopoly firm, it has little justification for regulating competitive markets. Regulating the behavior of monopolistic telephone companies as they compete in external markets has proven to be one of the central problems of telecommunications policy in the 20th century.

As the capabilities and the uses of the public network began to extend to non-telephony matters, regulators and policy makers have struggled to cope with this expansion of the role of the network. The history of telecommunications policy reflects a general view of the network as a universal, monolithic, public network. The broadening of the uses, services, and capabilities of the network has, for the most part, been treated not so much as a change in the role of the network as a departure from the basic functions of the public switched telephone network. The difficulty with this approach to policy is that, by treating new services as extensions to the basic network, it fails to take into account the transformation of the basic network itself. This difficulty is most evident in the various attempts by the Federal Communications Commission to wrestle with the growing convergence of the computer and telecommunications industries.

To understand this traditional approach to policy, it is useful first to examine its origins. In particular, one must start with the development of the integrated, ubiquitous telephone system known as the public switched telephone network, or PSTN.

3.2 DEFINING THE PUBLIC TELEPHONE NETWORK

The modern PSTN, the network that enables one to connect to virtually anyone in the world by simply dialing a telephone number, owes its form, to a large extent, to the vision of former AT&T President Theodore Vail. Vail, who ran AT&T from 1907 to 1919, recognized the network externalities inherent in a telephone system. The value of the service, to each customer, rises with the number of customers connected to the network. The strength of the Bell System, he argued, lay in its universality—at the time, it could reach farther than any other telephone system. Vail's vision of the future of telephone service was summarized by his oft-quoted credo: "one policy, one system, one universal service" (Coon, 1939).

At the commencement of Vail's term, AT&T did not enjoy the monopoly that was to characterize it through most of the 20th century. Nearly half of the telephone lines in the country were operated by independent telephone companies, some in areas not served by AT&T and others in direct competition. Vail disdained competition in the provision of telephone service, and frequently railed against the wasteful duplication of facilities. To achieve his vision of an integrated, universal system, he pursued a strategy of both crushing the independents—by refusing them interconnection with AT&T's inter-city lines—and acquiring them. These policies were ultimately derailed when Vail's desire for integration overstepped the bounds of the contemporary antitrust climate. Believing that there were synergies in the combination of telephone and telegraph networks, Vail acquired Western Union and, in so doing, prompted an antitrust action by the Department of Justice.

3.2.1 The Kingsbury Commitment

The antitrust suit was settled out of court in 1913 in an arrangement that has come to be known as the Kingsbury Commitment, after the AT&T Vice President Nathan Kingsbury. As part of the settlement, AT&T agreed to three primary terms: 1) it would divest itself of Western Union, 2) it would no longer acquire competing independent telephone companies,[10] and 3) it would allow all competing independents to connect to its long-distance facilities. The settlement was prompted by AT&T's recognition of its precarious political position as an increasingly large, monopolistic corporation. Bornholz and Evans (1983) note that at the time of the antitrust action, the Interstate Commerce Commission, having recently been given jurisdiction over telecommunications by the 1910 Mann-Elkins Act, had begun an investigation of AT&T. In addition, the Postmaster General had been advocating nationalization of the telephone system. The time, it would appear, was ripe for compromise.

Implicit in the Kingsbury Commitment, as Temin (1987) suggests, was AT&T's long-term acceptance of regulation. In return, its monopoly over much of the nation's telephone service would be sanctioned by the government. This bargain—monopoly status at the price of government regulation—has been the hallmark of telecommunications policy ever since. The Kingsbury Commitment, by mandating the interconnection of competing independents to AT&T's "Bell System," paved the way for the unified system envisioned by Vail. From 1912 to 1921, the percentage of telephones operated by independents not connected to the Bell System

---

[10] AT&T was allowed to continue acquiring independents who operated in territories not served by AT&T.

dropped from 17% to 4%.[11] During the same period, the percentage of telephones operated by independents also declined, from 45% to 36%.[12]

The Bell System continued to take shape during the 1920s, spurred by the Willis-Graham Act of 1921, which lifted the prohibition on the acquisition of competing independent telephone companies. By the time of the passage of the Communications Act in 1934, the number of telephones operated by independents made up only 21% of the network and the percentage of telephones not connected to the Bell System was negligible.[13] The goal of a unified, nationwide telephone network had essentially been reached.

3.2.2 The Communications Act of 1934

The view of the telephone network as an integrated whole is embodied in the 1934 Communications Act, which created the Federal Communications Commission. The Commission, which inherited the responsibilities of the Interstate Commerce Commission for regulating the telephone industry, was created

> "for the purpose of regulating interstate and foreign commerce in communication by wire and radio so as to make available, so far as possible, to all the people of the United States a rapid, efficient, Nation-wide, and world-wide wire and radio communication service."[14]

The purpose of the FCC was not simply to check the power of the telephone monopoly. In addition to traditional regulatory supervision, the FCC had a responsibility to ensure the proper functioning of the telephone network. The network had become a valuable public resource whose preservation warranted federal attention.

Although containing little new in the way of substance—much of the language governing the regulation of the telephone companies was taken directly from the Interstate Commerce

---

[11] Federal Communications Commission (1938). *Telephone Investigation: Proposed Report.* Government Printing Office, Washington, DC, hereafter referred to as *Walker Report*, cited in Bornholz and Evans (1983).

[12] Ibid.

[13] Ibid.

[14] Communications Act of 1934 § 1, 47 U.S.C. 151.

Act—the Communications Act did reaffirm two important principles that continue to shape our perception of the network. The first is the principle of common carriage, or nondiscrimination. Telephone companies were classified as common carriers, whose responsibilities are defined in Title II of the Act. Common carriers are not to deny any reasonable requests for services that they are engaged in providing. Furthermore, it is unlawful for a common carrier to make any "unjust or unreasonable discrimination in charges, practices, classifications, regulations, facilities, or services."[15] The second key principle emanating from the Communications Act is that of universal service, mandated by the charge, cited above, to make the communication service available "to all the people of the United States."

By 1934, the public telephone network had taken shape. It was a single, nationwide, interconnected system, open to all without discrimination. Service was to be extended to all of the nation's citizens. Although it consisted of the interconnected networks of many carriers, the network was dominated by a single firm, AT&T, who held primary responsibility for its operation as a whole. Viewed by the states and the federal government as a natural monopoly, it was regulated in lieu of competition.

3.3 COMPUTERS AND THE INTRODUCTION OF NON-VOICE APPLICATIONS

An additional characteristic of the public network, as it developed, was the service it provided. It was a telephone network. The Bell System's responsibility for the network extended all the way through the telephone instrument itself. In this arrangement, the network provider had control over the use of the network because it manufactured, owned, and maintained the terminal equipment. The advent of digital computers and the imagination of communications users led to alternative uses of the telephone network, as described in Chapter 2. These

---

[15] 47 U.S.C. 202.

developments posed difficult questions for the regulators of the Bell System—as new applications of the regulated network are developed, should they too be regulated?

3.3.1 The 1956 Consent Decree

Policy solutions to the question of how to treat novel uses of the public network were constrained for many years by restrictions on the lines of business open to AT&T. Shortly after its inception, the FCC began an extensive investigation into the business practices of AT&T. The investigation[16] suggested that AT&T was purchasing equipment from its sole source affiliate, Western Electric, at inflated prices. The cost of this equipment was passed on to the consumers of telephone service in the form of higher rates. World War II postponed any action on the report's allegations until 1949, when the Justice Department filed an antitrust suit against AT&T. The suit was settled in 1956 with a consent decree that required Western Electric to license its patents and forbade Western from manufacturing equipment unrelated to the provision of common carrier services.[17] The most significant provision of the decree, however, was a limitation on the businesses in which AT&T could engage. AT&T was enjoined from participating in "any business other than the furnishing of common carrier communications services." This restriction, which effectively barred AT&T from the computer and data processing industries, limited regulatory responses to the growing demand for computer communications.

3.3.2 The Carterfone Decision

As long as AT&T held a monopoly on terminal equipment—as long as it could decide what was to be connected to the network—it controlled the use of the network. This control, along with the restrictions mandated by the 1956 Consent Decree, was an obstacle in the path of innovation.[18] AT&T's discretion over equipment connection discouraged the development of

---

[16] Walker Report, *op. cit.*

[17] *United States v. Western Electric Co.*, 1956 Trade Cas. 71, 134, (D.N.J. 1956).

[18] Although Bell Laboratories is famous for its outstanding record of technological innovation, these innovations were not, by and large, translated swiftly into new services (Bolter and McConnaughey, 1991).

new terminal equipment by third parties and the consent decree prohibition on unregulated business activities prevented AT&T from venturing too far beyond the basic application of telephony. The issue of control over terminal equipment, and hence control over the use of the network, came to a head with the FCC's landmark 1968 *Carterfone* decision.[19]

The Carterfone, manufactured by Carter Electronics, was a device that patched mobile radio telephones into the PSTN, not by direct electrical connection, but by acoustic coupling. Several of AT&T's Bell operating companies, whose interconnection tariffs prohibited the connection of third-party devices to the network, disconnected the service of a number of Carter's customers. When Carter sued AT&T on antitrust grounds, the case was remanded to the FCC, which refused to accept AT&T's argument that, in order to carry out their responsibility to "establish, operate and improve the telephone system," the telephone companies must have "absolute control over the quality, installation, and maintenance of all parts of the system."[20] The FCC accordingly ordered the interconnection tariffs invalid. AT&T responded by developing a general solution for the interconnection of third-party equipment. It would allow interconnection mediated by what it termed "protective coupling arrangements" (PCAs), designed and manufactured by AT&T to buffer the network from the possible harmful effects of foreign equipment. Although the PCA solution proved problematic and was dropped in 1975 in favor of an FCC equipment registration program, the move marked the beginning of the competitive era in terminal equipment. AT&T had defined the limit of its responsibility (or hegemony, depending on one's point of view), effectively ceding control over the applications of the network to the users. *Carterfone* redefined the public network—it was not just for telephone and other applications selected by the Bell System—its applications were ultimately under the control of its customers.

---

[19] In the Matter of Use of the Carterfone Device in Message Toll Telephone Service, *Decision*, 13 FCC 2d 420 (1968).

[20] 13 FCC 2d at 424.

3.3.3 The FCC Computer Inquiries

The Carterfone case arose during the first of what was to become a series of FCC inquiries into the policy ramifications of the computer industry and its communications applications. The Commission had a responsibility to regulate common carrier communication services under Title II of the Communications Act. The data processing industry, on the other hand, appeared to need no regulation—it was considered to be highly competitive. The overlap between the two industries led the FCC to launch an inquiry into the policy problems associated with the "interdependence of computer and communication services and facilities" in 1966.[21]

The First Computer Inquiry

The first exploration into the relationship between computer and communications services led to an attempt to distinguish data processing services from the communications services to be regulated by the Commission under the Communications Act. The data processing industry, the Commission concluded, was competitive and flourishing—it was to be left unregulated. The Commission did, however, recognize the existence of a gray area between communications and data processing. It defined a "hybrid" service as the "offering of service which combines Remote Access data processing and message-switching to form a single integrated service."[22] Hybrid services were to be further categorized into hybrid communication services and hybrid data processing services, depending on which function, the message switching or the data processing, was predominant. Hybrid communication services could be provided by common carriers as regulated, tariffed offerings. The Commission interpreted the 1956 Consent Decree to prohibit AT&T and its affiliated Bell Operating Companies (the BOCs)

---

[21] In the Matter of Regulatory and Policy Problems Presented by the Interdependence of Computer and Communication Services and Facilities, (hereafter First Computer Inquiry) *Notice of Inquiry*, 7 FCC 2d 11 (1967).

[22] First Computer Inquiry *Order*, 40 FCC 2d 293 (1973) at 295. The term "message-switching," as opposed to circuit-switching, referred to the transmission of "a customer's set message at a charge based upon the information sent" (First Computer Inquiry, *Tentative Decision of the Commission*, 28 FCC 2d 291 (1970) at 296). Message switching, the dominant paradigm of the telegraph industry, involved a store-and-forward function akin to today's electronic mail or voice messaging services.

from engaging in data processing services or hybrid data processing services. Other common carriers were allowed to provide these unregulated services, but only from structurally separate subsidiaries.

The Second Computer Inquiry

With the benefit of hindsight, one can see the potential for failure in the *Computer I* service classifications. If the concern over the growing data processing industry was the potential for interdependence with communication services, one can imagine that a growing number of services would thus fall into the "hybrid" category. The convergence was hastened by the development of distributed computing. Computing capabilities were no longer centralized in remote hosts, but were inherent in user terminals as well. As terminals began to perform data processing functions that were more than incidental to their communications functions, the distinction between hybrid communication and hybrid data processing services became impractical. This complication was made all the more apparent when AT&T proposed a tariff for an intelligent data communications terminal, the Dataspeed 40/4, that offered data storage, retrieval, and editing capabilities. The Dataspeed terminal, as it was furnished by the dominant common carrier, AT&T, clearly had to be regulated,[23] but, insofar as it offered data processing functions, appeared to fall in the unregulated category of data processing. Recognizing that the *Computer I* framework had been rendered obsolete by the changes in the nature of computing, the Commission launched the Second Computer Inquiry, beginning in 1976.[24]

In their second attempt to address the issue, the Commission accepted the inevitability of the convergence of computers and communications. Noting that it was impossible to draw "an

---

[23] While the *Carterfone* decision prohibited AT&T and its affiliates from refusing to interconnect third-party equipment without cause, it did not deregulate the terminal equipment market. Common carriers furnishing terminal equipment continued to do so under tariff until the final decision of the Second Computer Inquiry, which deregulated the market for customer premises equipment.

[24] In the Matter of Section 64.702 of the Commission's Rules and Regulations (Second Computer Inquiry), *Notice of Inquiry and Proposed Rulemaking*, 61 FCC 2d 103 (1976).

enduring demarcation" between data processing and communication services, the Commission adopted a new set of classifications.[25] This new approach reflected the changed nature of the public network, since the Carterfone decision had removed the limitations on its use. Services were classified into a dichotomy of "basic" and "enhanced" services. Basic services were limited to "the common carrier offering of transmission capacity for the movement of information." The Commission recognized that the provision of basic services could contain elements of data processing and allowed that "data processing, computer memory or storage, and switching techniques can be components of a basic service if they are used solely to facilitate the movement of information." The strict definition of enhanced services was somewhat obscure—"any offering over the telecommunications network which is more than a basic transmission service." The Commission did, however, subsequently expound upon this new classification, stating that

> "the term enhanced service shall refer to services offered over common carrier transmission facilities which employ computer processing applications that act on the format, content, code, protocol or similar aspects of the subscribers' transmitted information; provide the subscriber additional information, or involve subscriber interaction with stored data."[26]

All common carriers, including AT&T and the BOCs, were allowed to create structurally separate subsidiaries to engage in the provision of enhanced services. AT&T's freedom to venture into the area of unregulated services seemingly forbidden by the 1956 Consent Decree was justified by an assessment that any unregulated operations of AT&T would be incidental to their provision of common carrier services.

The *Computer II* decision redefined the public network yet again from a policy standpoint. What was "basic" to the network was not so much the provision of telephone service but the "common carrier offering of transmission capacity." The telephone network was thus, in the eyes of its federal regulators, a communications network first and foremost. Although the basic role of the network was no longer limited to telephone service, the *Computer II*

---

[25] Second Computer Inquiry, *Final Decision*, 77 FCC 2d 384 (1980).

[26] Section 64.702(a) of the Federal Communications Commission's Rules, 47 CFR § 64.702(a).

classifications do not reflect a new approach to telephony. Telephone service remains basic, in *Computer II* and later rulings, insofar as it can be viewed as the provision of transmission capacity, on a switched basis, between the caller and the called. It is thereby subsumed in the definition of the basic role of the network. Thus, while the basic functions of the network had been generalized in application-independent terms, the primary traditional application of the network remained one of its basic functions.

The Third Computer Inquiry

The FCC recognized that some services may be difficult to classify under the *Computer II* distinctions and that services, combining basic and enhanced functions, may be most efficiently provided on an integrated basis. Accordingly, the Commission allowed for the network providers to request waivers of the structural separation requirement if they could show that the benefits of vertically integrated provision of a service outweighed the costs associated with the potential for cross-subsidization or other forms of anti-competitive behavior. Having received a number of waiver requests, including those in which AT&T or the BOCs would argue that a service could not be provided except on an integrated basis, the Commission opened the Third Computer Inquiry in 1985 to re-examine the merits of the *Computer II* framework (Smith and Pitt, 1991).

The Third Computer Inquiry began against the backdrop of a much changed telecommunications industry. The 1984 AT&T Divestiture, the result of an eight-year antitrust litigation, broke up the Bell System. AT&T retained it long distance services, manufacturing, and research operations, but relinquished its local affiliates, the Bell Operating Companies. The settlement, in the form of the Modification of Final Judgment, or MFJ, freed AT&T to pursue business ventures in the computer and data processing industries.[27] The BOCs, on the other hand, were limited in their outside ventures. They were prohibited from engaging in long-

---

[27] *United States v. AT&T*, 552 F. Supp. 131 (1982)

distance services, manufacturing, or the provision of "information services."[28] The restrictions on the newly independent BOCs were justified by their control of the "bottleneck" of local exchange services. AT&T could be freed from its primary restrictions because it was to be engaged in competitive markets; the BOCs, with monopoly power over bottleneck facilities, had incentives and, presumably, opportunities to compete unfairly in unregulated markets.

The FCC ruled, in its *Computer III* decision,[29] that the economic costs of the structural separations mandated by *Computer II* outweighed their benefits in checking the monopoly power of AT&T and the BOCs. The basic/enhanced dichotomy of *Computer II* was retained, despite the Commission's concerns about its feasibility,[30] but the structural separation of providers was lifted in favor of non-structural safeguards. AT&T and the BOCs could provide enhanced services and basic services on a integrated basis, subject to the Comparably Efficient Interconnection (CEI) and Open Network Architecture (ONA) requirements outlined in *Computer III*. ONA was the Commission's long-term solution. It directed AT&T and the BOCs to offer basic services on an unbundled basis. The tariffed, unbundled services would be used by the enhanced services operations of the network providers and the independent enhanced service providers (ESPs) to construct enhanced services. In principle, so long as the monopoly carriers and the ESPs were working from the same menu of features and services, subject to the same

---

[28] Information services were defined in the MFJ to include

> "the offering of a capability for generating, acquiring, storing, transforming, processing, receiving, utilizing, or making available information which may be conveyed by telecommunications, except that such service does not include any use of such capability for the management, control, or operation of a telecommunications system or the management of a telecommunications service." *United States v. AT&T*, 552 F. Supp. 131 (1982) at 229.

[29] In the Matters of: Amendment of Sections 64.702 of the Commission's Rules and Regulations (Third Computer Inquiry); and Policy Concerning Rates for Competitive Common Carrier Services and Facilities Authorizations Thereof; Communications Protocols under Section 64.702 of the Commission's Rules and Regulations, *Report and Order*, 104 FCC 2d 958 (1986).

[30] The Commission, in a Notice of Proposed Rulemaking, had suggested the creation of a third category, "services ancillary to communications," that would be regulated or unregulated depending on the particular economics of the service. Concern over the potential for regulatory uncertainty in this approach led the Commission to abandon it in favor of other non-structural safeguards.

rates and terms, competition in enhanced services would take place on an equal basis. CEI was to be an interim step towards ONA. If the BOCs or AT&T wished to offer an enhanced service before their ONA filings had been accepted, they would be required to file a CEI tariff demonstrating that other service providers would have comparable access to the features and facilities used in the provision of the service.

The *Computer III* approach can be seen as a victory for AT&T and its orphaned affiliates, the BOCs, in that they gained the freedom to offer enhanced and basic services in a vertically integrated manner. This freedom, however, came at the expense of subjecting themselves to unbundling requirements. While the FCC left the details of the unbundling required by ONA to industry fora, *Computer III* did set forth a principle that certain functions performed in the network cannot be fully integrated. The principle of unbundling suggests that the network can be broken down into functional components and that these components must be offered independently, with well-defined interfaces. Whether or not the FCC intended a fundamental unbundling of the network and a passing of control of many aspects of the operation of the network from the network providers to the enhanced service providers, it initiated a process that could achieve that result. By mandating an "open network," the Commission sowed the seeds for disputes over what constitutes an adequate degree of openness and thereby put pressure on the network providers to relinquish much of their control over the network.

3.4 SUMMARY

The public telephone network grew up as a single service network. Telephony was the only application of this network. The independent telephone companies notwithstanding, the network was, for all intents and purposes, owned, operated, and maintained by a single, vertically integrated, monopoly provider. The network provider controlled all aspects of the network, from the telephone instruments to the manufacturing of the network equipment. The invention of the digital computer and the new forms of communications it spawned led to changes in the use of the public telephone network. Federal policy, in wrestling with these changes, effectively redefined the public network to the point where it is now a multi-service network. Policy decisions have also broken the monopoly provider's control over many aspects of the network. Terminal equipment, and by extension the applications of the network, is now completely deregulated. Open Network Architecture has the potential to transfer significant amounts of the intelligence of the network away from the network provider.

As the concept of the single, vertically integrated network provider crumbled, policy makers recognized that while some network functions and services appeared to have the characteristics of natural monopoly, others could be better served by competition. Regulating the participation of the monopoly provider in these competitive markets has led the FCC to classify network services as either basic (regulated) or enhanced (unregulated). Implicit in the definitions of these categories is a recognition of the role of the network as a provider of generic communications capacity. The term "basic service" does not refer specifically to telephone service—it refers, in the language of the *Computer II* decision, to the "offering of transmission capacity for the movement of *information*." Telephone service is, however, considered a basic service.

The following chapters, on the CCITT's plans for Broadband ISDN and Asynchronous Transfer Mode, the technology that will be employed to realize those plans, will demonstrate how the public network of the future can be conceptualized as a basic platform network supporting many services and applications. Telephone service, in this view, can be understood as one of many applications of this basic platform network, rather than as a basic service. The analysis of B-ISDN and ATM will also reveal the potential for creating open architecture in future public networks.

Chapter Four

Broadband ISDN and Asynchronous Transfer Mode

4.1 INTRODUCTION

When the CCITT began its development of Broadband ISDN, it was envisioned as an extension of the ISDN principle developed in Narrowband ISDN to facilities of greater bandwidth, or higher bit rate. While N-ISDN was designed to support a broad range of communication services, it was based on the existing infrastructure of copper plant, and thus limited in the data rates it could accommodate. N-ISDN offers a basic interface of 64 Kbps B channels and cannot support the high-speed data transfer applications (which, for now, run on the order of 10-100 Mbps) found in local area networks, nor can it deliver standard television quality video signals, that, with current compression techniques, require approximately 1.5 Mbps. N-ISDN is thus unable to integrate more than a limited range of communication services, namely voice and some low-speed data applications. Broadband ISDN, because it assumes transmission over optical fibers, would be capable of data rates at least a thousand times greater than N-ISDN. With such ability to support very high data rates, B-ISDN, in principle, would be capable of integrating all types of communications applications, ranging from telemetry up to high-definition television and beyond.

4.2 BROADBAND ISDN SERVICES

In its first published recommendation on Broadband ISDN, the CCITT specified the services for which Broadband ISDN would be designed.[31] These services break down into two categories: interactive services and distribution services. Interactive services are subdivided into the following groups.

> *Conversational Services* are defined as those with real-time, bi-directional, end-to-end transfer of information. Examples of conversational services include traditional telephony and videotelephony.
>
> *Messaging Services* involve non-real-time, user-to-user communications that are facilitated by storage units that perform store-and-forward and mailbox functions. Examples include voice messaging, electronic mail, and future broadband services such as video or multimedia mail.
>
> *Retrieval Services* are those in which individual users can, on demand, retrieve information stored in information centers. The information can be retrieved on an individual basis and the user can control the start of any information sequence. Examples of retrieval services would include legal or news information archives (such as Lexis/Nexis), or databases that contain photographic images or video clips.

---

[31] The discussions in this chapter, both of the CCITT's recommendations for Broadband ISDN and of the principles of Asynchronous Transfer Mode, are based on De Prycker (1991) and CCITT Recommendations I.121, I.150, I.211, I.311, I.327, I.413, and I.610.

The distribution services refer to those in which a central source provides a continuous flow of information to a large number of users. Broadcast radio and television are examples of distribution services, as are services that provide continuous data on current stock market prices.

4.3 TECHNICAL OPTIONS: CIRCUIT SWITCHING OR PACKET SWITCHING

The challenge for the CCITT was to design a single network capable of supporting such a broad spectrum of services. The services envisioned for B-ISDN not only encompassed a wide range of data rates, they varied considerably in two other important respects. First, some applications require continuous streams of data while others are bursty in nature. Voice, for example, is typically coded using Pulse Code Modulation and produces a continuous, 64 Kbps stream of data, whereas interactive data processing is characterized by the transmission of bursts of data corresponding to user input and system response. Second, some services, such as voice and many video applications, require end-to-end synchronization and low delay through the network. Whereas a data file transfer or a database query can withstand some network delay—the only effect is a perceived sluggishness of the system—a voice telephone call or a live video program suffers greatly if its transmissions arrive sporadically. Imagine watching a television program where the images slow down and speed up randomly! Any solution for B-ISDN has to accommodate services and applications with these widely varying characteristics and thus the flexibility of the network becomes paramount.

The dominant paradigm of the telephone network and, consequently, Narrowband ISDN, is circuit switching. Circuit switching is especially well-suited to applications that produce continuous bit streams and that require end-to-end timing. However, as noted in Chapter 2, one disadvantage of circuit switching is its inherent inefficiency in supporting bursty applications. A circuit is held open throughout a conversation and thus, during periods of relative inactivity or silence during conversations, as in the case of a database query or a telemetry application, network resources are wasted. A second weakness of circuit switching is the necessity to define channel rates. If a circuit is to provide end-to-end transmission between users, it must operate at a particular data rate. Users are thus limited to exchanging information at rates no greater than the capacity of the channel and, if they choose to converse at lower speeds, the network is providing excess capacity. One solution considered for B-ISDN was to define a series of different standard channel rates that would be switched. Users could then tailor their applications to meet these specifications. However, such a solution does not solve the problem of inefficiencies due to burstiness. A further concern with standardizing channel rates is the fundamental lack of knowledge of specific requirements of future services and applications. Choosing appropriate rates would thus involve a certain amount of guesswork.

Packet switching is the most obvious alternative to circuit switching.[32] Packet switching has efficiency advantages for bursty applications in that it does not require the provision of an open circuit, but is disadvantageous in applications that require end-to-end timing. The delays inherent in packet switching, produced by processing information into data packets and occurring as the result of network congestion, make conventional packet switching techniques unsuitable for voice applications.

4.4 THE SOLUTION: ASYNCHRONOUS TRANSFER MODE

The solution at which the CCITT arrived is based on a technique known generally as fast packet switching. Fast packet switching is a stripped-down version of traditional packet switching. By limiting the functions, and thus reducing the complexity of the packet switch, the very high transfer speeds needed for broadband switching can be achieved.

Fast packet switching differs from its ancestors by employing significantly shorter data packets. One delay in packet switching is associated with the time required to segment information into data packets at one end of the network and reassemble those packets into usable information at the other. One can imagine that the longer a packet, the longer it takes to

---

[32] See section 2.3.2 for an overview of packet switching principles.

assemble, thus, in order to minimize this assembly/reassembly delay, short packets are desirable. Reducing the packet size also reduces the buffering requirements at each node, further simplifying switch design.

Another simplification is a reduction in error handling capabilities. Traditional packet switching protocols typically perform error correction at each node of the network. Fast packet switching limits this node-by-node error correction to only the header section of each data packet, and leaves additional error handling to the applications themselves, to be performed on an end-to-end basis.[33] This change in philosophy was made possible by advances in transmission systems, namely the use of optical fibers and digital coding, which significantly reduce the probability of transmission errors. Furthermore, different applications have different tolerances for communication errors. For example, a telephone conversation is not noticeably degraded by a handful of single bit errors, whereas an errant bit in an electronic funds transfer could be catastrophic. Given this range of tolerance for errors, it is appropriate that the error handling functions take place in the applications rather than in the network.

For Broadband ISDN, the CCITT has chosen to support all applications, whether high-speed or low, with continuous or bursty bit streams, with a fast packet switching technique known as Asynchronous Transfer Mode.

4.4.1 The ATM Cell

The fundamental unit of the digital public telephone network is the 64 Kbps channel, which carries exactly one voice conversation. All switching in the network (excluding that done by peripheral switching services such as packet-switched or Frame Relay services) is done on a channel by channel basis. All transmission is done in groups of 64 Kbps channels—24 channels

---

[33] In any packet switching technique, each data packet consists of a header, which includes routing information and other information necessary to the functions of the network, and the payload, or the data being transmitted through the network. Error correction on the header information is vital because any errors in the header can lead to misdelivered packets, which, unlike errors within the information payload, cannot be easily corrected at the communication endpoints.

make a 1.5 Mbps T1 (or DS-1) circuit, 28 DS-1 bundles are grouped into 45 Mbps DS-3 bundles. The fundamental unit of ATM is the 53-byte cell.[34] All information, whether voice, video, or computer data, exchanged over the network will be segmented into 48-byte chunks, to which 5-byte headers are appended. These cells are routed over the network, cell by cell, on the basis of their header information—no dedicated channels are assigned to a connection.[35]

One consequence of the cell by cell switching approach is the rate-independence of the system. The interface at which a user will be connected to the network will be defined for a particular data rate.[36] This access channel can be thought of as a conveyor belt, onto which many boxes can be placed. Just as one can place as many or as few boxes on the conveyor belt per unit time as one wishes (subject to the limit that no two boxes can occupy the same space), an ATM terminal can insert data cells into the access channel bit stream very rapidly, rather infrequently, or at any rate in between (Figure 4.1). Thus, for example, 64 Kbps voice information would generate a few hundred cells per second, whereas a high-resolution video signal producing an average of 90 Mbps of information would generate several hundred thousand cells each second. Once in the network, the data rate of the application is unimportant—the cells are forwarded to their target destinations individually—so long as the application is not generating cells faster than the network can process them.

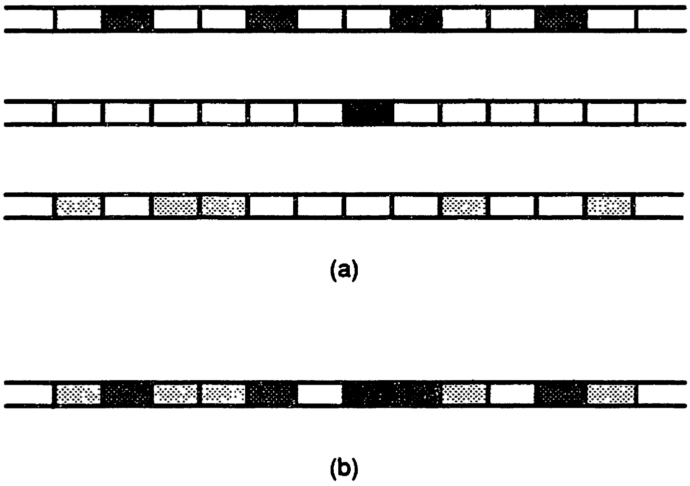
---

[34] An ATM data packet is called a "cell" to distinguish it from the packets used in traditional packet switching.

[35] The term "dedicated" is used to refer to communication paths that, once established between two endpoints, are available only to the users associated with those endpoints. A leased, private transmission line is an example of a dedicated channel, as is the path connecting two parties in a circuit-switched telephone call. Dedicated channels are thus distinguished from virtual channels, in which cells or packets travel from one endpoint to another over shared communication channels.

[36] The CCITT has, for the moment, defined user interfaces operating at 155.52 Mbps and 622.08 Mbps. However, the ATM switching principle is scalable—it can be extended to support higher speeds.

Figure 4.1.
The ATM user access arrangement.



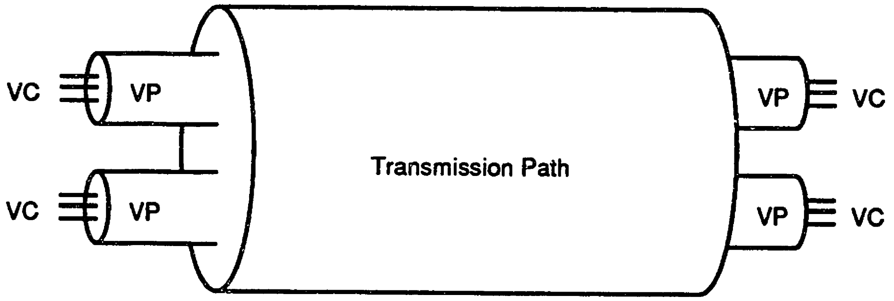
ATM cells associated with three separate communications applications (a) are multiplexed together in the user access channel (b). The access channel provides a synchronous data stream into which cells can be inserted by individual applications according to the frequency at which they are generated.

4.4.2 Cell Routing

ATM is a connection-oriented transport mechanism. As in a telephone network, "calls" are set up between terminals (or end users) for the exchange of information and released when the exchange is completed. However, in an ATM environment, these connections are not dedicated paths, but rather virtual connections. When an ATM user initiates a call to another ATM user, the network assigns a virtual connection between the two users. This virtual connection represents a designated routing path through the network. Every cell generated by the user travels through the network along this particular path. A connection-oriented service is distinguished from a connectionless service, in which each cell contains explicit routing instructions, such as the full destination address. The United States Postal Service is an example of a connectionless service—the destination address is written on each envelope that is mailed.

The header portion of every ATM cell contains the virtual connection information in the form of a two-part identifier. This label consists of a virtual channel identifier (VCI) and a virtual path identifier (VPI). The VCI refers to an individual service component of a particular user-to-user connection. For example, in a video telephone call between two people, separate VCIs could identify the audio, the video, and perhaps any text information that is exchanged between the two users. A virtual path is a bundle of virtual channels (Figure 4.2). For example, a local telephone company connects its central office switches with trunks that carry calls linking the subscribers of different exchanges. If these calls were to be transmitted as ATM cells, all cells, representing all of the telephone calls between the two exchanges, might carry the same virtual path identifier, but the VCI would link each cell with a particular call.

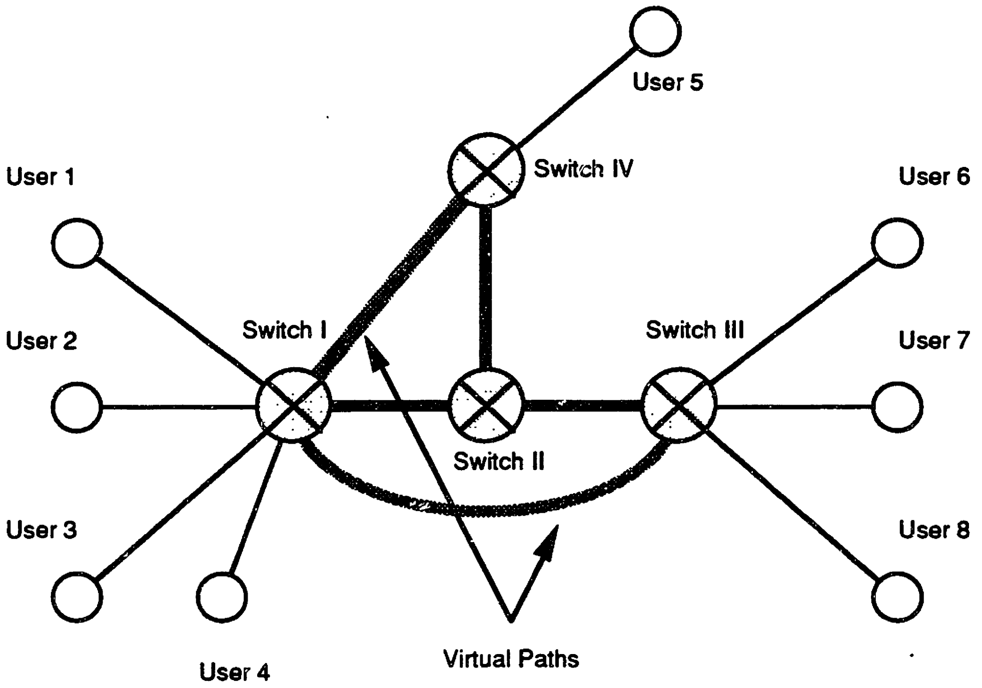
Figure 4.2.
Relationship between the virtual channel, the virtual path, and the transmission path.



Source: CCITT Recommendation I.311

A typical use of virtual paths and virtual channels is shown in Figure 4.3. In this example, Users 1, 2, and 3 initiate calls to Users 6, 7, and 8, respectively. Unique virtual channels are established to connect User 1 to User 6, User 2 to User 7, and User 3 to User 8. These three channels, may, however, share a virtual path between Switch I and Switch III. The connection between User 4 and User 5 is given a separate virtual path identifier, which specifies the path from Switch I to Switch IV. Both virtual paths travel the same physical path between Switch I and Switch II; Switch II is responsible for analyzing the incoming cell headers, and based on the virtual paths identified, route the calls to either Switch III or Switch IV. The virtual path concept enables the use of an intermediate switching device. Note that in this example, Switch II does not examine the VCI of any incoming cell, and thus need not be as complex as the other ATM switches. Such a simplified ATM switch is called a virtual path-only (VP-only) switch.[37]

Figure 4.3.
A typical use of virtual paths.



The dark lines represent physical paths connecting the ATM switches.

---

[37] Interconnected VP-only switches form what are frequently referred to as ATM "cross-connect" networks.

When an ATM connection is established, a unique route through the network is designated. This designation is expressed by the virtual path and virtual channel identifiers. However, because, at any given time, an astronomical number of connections will be established in the network, the VPIs and VCIs do not have global significance, i.e. for a telephone call between a person in New York and another in Washington, no single VPI/VCI label is established. Instead, the VPIs and VCIs are translated at each node of the ATM network. A cell originating in New York may have one VPI/VCI from the user terminal to the first ATM switch, where it may be bundled with other cells heading to the next network node in a common virtual path. The establishment of the virtual connection,[38] the network route, requires that each node be updated, on a call-by-call basis, with translation information. Such activity requires continuous interaction between the nodes of the network and the control network responsible for establishing, maintaining, and tearing down virtual connections.

An additional feature of the ATM routing scheme is the opportunity for semipermanent connections. Both virtual channels and virtual paths can be assigned for continuous use, just as dedicated telephone circuits are often used to connect multiple locations in a private network. A telemetry application such as remote sensing may be well-served by a semipermanent virtual channel between the sensing device and the central located processor. Semipermanent virtual paths may be appropriate in a corporate customer's private network. The virtual path may be used to carry a variety of different services between two customer-owned ATM switches.

4.4.3 The Broadband Transmission Infrastructure

While the CCITT was investigating broadband switching techniques, Bellcore was developing specifications for a high-speed optical transmission network infrastructure. The proliferation of optical transmission systems had created the need for a standardization of optical

---

[38] The term "virtual connection" is used to refer to either a virtual channel connection or a virtual path connection.

interfaces. Most transmission systems, while providing the standard U.S. electrical interfaces of T1 (1.5 Mbps) and DS-3 (45 Mbps), used proprietary optical techniques. Consequently, telephone trunks between two carriers had to be converted from optical to standard electrical signals at the boundary of the carriers' networks. Bellcore's solution to this problem of interconnection is a series of interface specifications known as the Synchronous Optical Network, or SONET, standards.[39]

SONET specifies optical interfaces of 51.84 Mbps, 155.52 Mbps, 622.08 Mbps, and 2.488 Gbps. Standardization of a 10 Gbps interface is also expected (Kaiser, 1991). Within these high-speed optical streams can be organized the current digital hierarchy of channels. For example, the 51.84 Mbps interface was chosen to carry a 45 Mbps DS-3 signal, with the remaining capacity used for management and other overhead information. Individual T1 channels may also be multiplexed directly in and out of the SONET transmission streams. SONET also features robust management capabilities. Networks can be designed as self-healing rings, greatly increasing the reliability of public transmission networks. Channels within the SONET backbones can be administered and configured remotely, allowing for rapid "installation" of fixed rate transmission circuits.

The effect of the SONET specifications is to make possible a standard, flexible, high-speed digital transmission network. This optical transmission network will form a basic level of network infrastructure, connecting the various switching resources of the network. The SONET infrastructure is not limited to internal network transport, however. SONET transmission will be used to interconnect the networks of service providers to the networks of the telephone companies and to connect the nodes of private networks. SONET may also be used to extend optical interfaces from the public network directly to end customers.

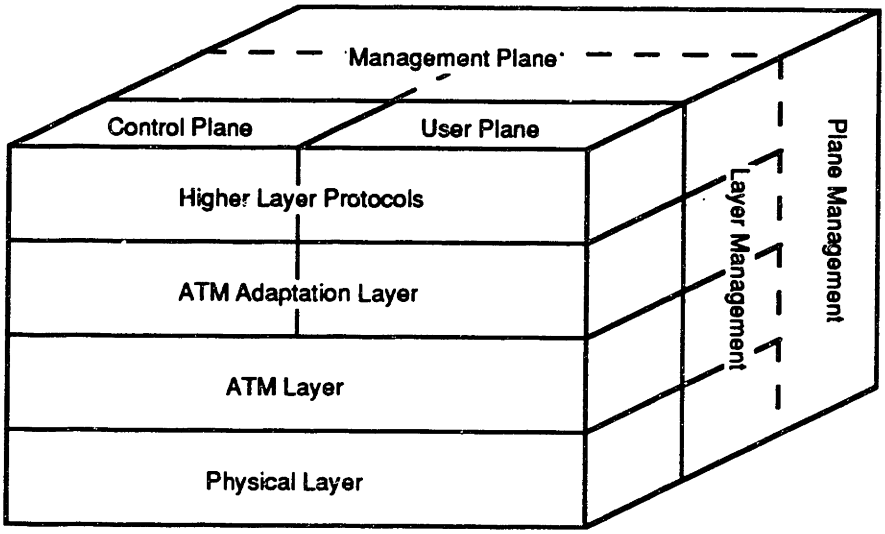
---

[39] For background on SONET specifications and features, see Sandesara *et al* (1990), Aprille (1990), and Ballart and Ching (1989).

4.4.3 The Broadband ISDN Protocol Reference Model

The B-ISDN Protocol Reference Model is shown in Figure 4.4. ATM networks will use a layered approach to protocols along the lines of the International Standards Organization's Open Standards Interconnection reference model.[40] Three protocol "planes" are defined, for carrying user, control, and network management information. The user and control planes share a common base of three CCITT-defined layers: the Physical layer, the ATM layer, and the ATM Adaptation layer.

Figure 4.4.
The Broadband ISDN Protocol Reference Model.



Source: De Prycker (1991)

---

[40] See Jain and Agrawala (1990) for an exhaustive treatment of the OSI model. The subject is also summarized in a number of data communications textbooks, including Stallings (1987) and Tanenbaum (1981).

The Physical Layer

The Physical layer is responsible for transporting ATM cells from node to node. To this end, the Physical layer packages ATM cells into the appropriate transmission vehicle, provides a mechanism for delineating cells, and performs error correction on the ATM header bytes. The Physical layer extends downward in the model to include the physical medium itself. At each node, the Physical layer takes in a stream of ATM cells and delivers an identical stream of cells to the next node in the network. For transmission between the nodes, the CCITT has recommended two options. The first is to transport ATM cells within SONET envelopes. While SONET is based on fixed rate channels and was not conceived with ATM in mind, ATM cells can easily be packaged into SONET channels (De Prycker, 1991). This approach adds a degree of flexibility in network design in that SONET channels carrying ATM traffic can share the same fiber and the same transmission equipment as SONET channels carrying fixed bit rate circuits, such as private lines or telephone trunks. Thus a network planner may build a SONET-based transmission infrastructure that is independent of the type of traffic it carries—a separate transmission network for ATM is not required. The other option suggested by the CCITT is a pure ATM transmission scheme. In this arrangement, fixed rate channels are replaced by a high-speed, continuous stream of ATM cells. Cells from different users arrive at an ATM multiplexer at one speed and are multiplexed together for transport to the next node in the network at a higher speed. Both solutions have certain merits; however, it is now widely accepted in the industry that, at least initially, SONET frames will be used to transport ATM cells across the network (Aaron and Dècina, 1991).

The ATM Layer

Transporting each ATM cell from its point of origin to its intended destination is the task of the ATM layer. The ATM layer receives cells from the Physical layer and must interpret the VPI/VCI labels. Based on this routing information, the ATM layer, at each node, performs any necessary VPI/VCI translations and selects the appropriate output path for each cell. The ATM

Layer is also responsible for multiplexing cells associated with many different virtual connections onto single streams of ATM cells. At the terminating point of the connection, the ATM layer delivers the proper cells to the ATM Adaptation Layer.

The ATM Adaptation Layer

The ATM Adaptation Layer (AAL) is responsible for preparing the data generated by a communications application for transmission over an ATM network. The ATM Adaptation Layer is divided into two sublayers: segmentation and reassembly (SAR) and convergence. The SAR sublayer reformats application data into 53-byte ATM cells at one end of the connection and reassembles cells into meaningful data at the other. The convergence sublayer may, depending on the specific application, perform such tasks such as maintaining synchronization (for applications requiring end-to-end timing), handling of misdelivered or lost cells, and correcting transmission errors. Although the ATM Adaptation Layer is confined to the edges of the network, left to the user terminals or to the processors of service providers, the CCITT is proceeding with efforts to develops standards for the AAL. Rather than attempting to develop standard AAL interfaces for individual communications applications, the CCITT has defined four general classes of applications, and will develop standard AAL specifications for each class. These four classes of applications are shown schematically in Figure 4.5. The applications are distinguished by whether timing is required between the source and destination, the character of the generated bit stream (constant or variable), and the connection mode (connectionless or connection-oriented).[41] AAL Type 1 represents real-time connection-oriented services such as telephony, where the bit stream is constant. Fixed rate private line services, such as T1 circuits, can also be transported across an ATM network using AAL Type 1 adaptation. Type 2 includes services such as variable bit rate video, where a real-time, end-to-end connection is established,

---

[41] ATM, as defined by the CCITT, is a connection-oriented transfer mode. Nevertheless, connectionless services can be transported across an ATM network through the use of semipermanent virtual connections. An example of connectionless data service over an ATM network is shown in Chapter 5.

but the bit rate fluctuates. Bursty data services, where no timing between the source and destination is necessary, can be supported as connection-oriented (Type 3) or connectionless (Type 4) services. In addition to the CCITT's recommended AAL types, the computer industry has proposed a fifth type, known as the Simple and Efficient Adaptation Layer (SEAL) (Cheung, 1992). SEAL is a stripped-down, limited function version of types 3 and 4, and can be used to support both connection-oriented and connectionless data services. The simplicity of SEAL offers advantages in terms of reducing both overhead and implementation complexity.

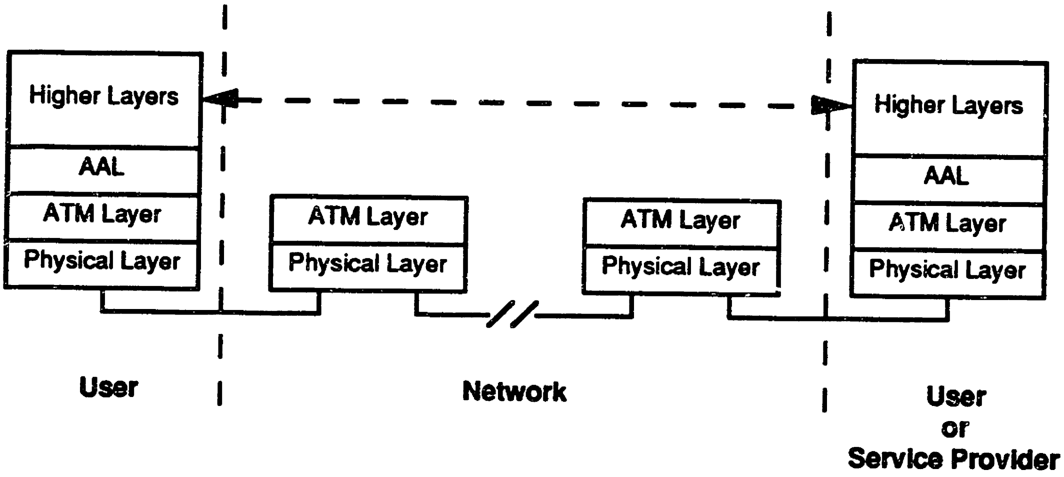
Figure 4.5.
Categorization of ATM Adaptation Layer types.

Type	1	2	3	4	5
Timing Between Source and Destination	Required		Not Required		
Bit Rate	Constant	Variable			
Connection Mode	Connection-Oriented			Connectionless	Connection-Oriented or Connectionless

Source: Cheung (1992) and De Prycker (1991)

The interaction of the three CCITT-defined protocol layers for B-ISDN is summarized in Figure 4.6. The Physical layer carries cells between nodes of the network, where they are processed by the ATM layer. ATM Adaptation Layer functions are not performed within the network, but at the edges of the network. The higher level protocols, which may include internetworking protocols or application-specific protocols, communicate from end-to-end.

Figure 4.6.
End-to-end communication between higher layer protocols.



4.4.4 Signaling and Control

The switching of ATM cells is quite simple once the virtual connections have been established—each switching node has only to read and translate the VPI/VCI of each incoming cell and route it to the appropriate output port. Establishing the virtual connection, or in many cases, connections, associated with a particular call is more complicated—it requires signaling negotiations between the user and the network, and within the network itself. In traditional telephony, signaling took place within the voice band; the pulses or tones used to signal the network were carried over the same channel as the caller's voice. In recent years, however, demand for sophisticated calling services has required greater signaling complexity. To allow for more flexibility, signaling messages are now carried over an overlay data network known as the System Signaling 7 (SS7) network. The establishment of a signaling network separate from the voice network is the first step in the implementation of a long-term approach to network control known as the Intelligent Network.

The Intelligent Network

The Intelligent Network approach to signaling and control is based on a migration of the network intelligence from local central office switches to remote, centralized processors.[42] In effect, it separates the logic of the switching function from the physical switching of circuits. This separation is a sharp departure from the traditional approach to switching, in which both functions were performed by local switches.

In 1965, AT&T introduced the first electronic switching system, the 1ESS, which utilized a technique known as stored program control. With the introduction of stored program control, the telephone switch began to resemble a computer—certain switch functions, such as the interpretation and translation of digits, were performed according to instructions from computer programs running inside the switch. The computer program approach proved highly flexible and made possible the multitude of custom calling features, such as call waiting, call forwarding, and abbreviated dialing, that are available today. Central office switches grew more and more complex, offering an ever increasing variety of software controlled features and services. The result of this development was a decentralized network architecture—the intelligence of the network was found in each of thousands of local central office switches.

The decentralized approach to network intelligence presents significant drawbacks. First and foremost is the problem of service deployment. If a telephone company wishes to deploy a new service, it must update the software of each central office switch individually, at considerable expense and logistical inconvenience. This problem is further complicated by the variety of switches typically found in a carrier's network. Not every switch runs the same operating software and thus, even if each switch is upgraded to provide the new service, it may operate differently in different service areas. Another key drawback for the telephone companies is that the software to control their central office switches is bundled with the switch hardware.

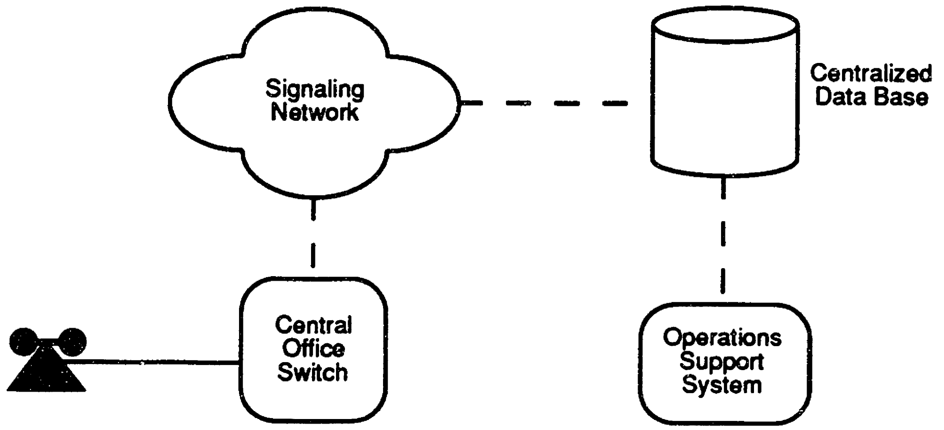
---

[42] For an overview of Intelligent Network principles and examples of its implementations, see Robrock (1991).

Because the hardware and software of a central office switch are provided as a package deal, the telephone company is dependent on the switch vendor for feature and service enhancements. Given that the central office switch market in the United States is effectively a duopoly (Hack and Downs, 1991), the telephone companies have little control over new service development.

The Intelligent Network is the industry's solution to the problems of service innovation and deployment. The most significant Intelligent Network concept is the removal of much of the service logic, or the call processing intelligence, from the local central office switch to remote processors (see Figure 4.7). The switch no longer handles all call processing tasks. Instead, the switch, upon recognizing that a call needs special treatment, sends a query to a centralized data base. The data base, containing specific call processing information, responds with instructions to the switch. In the Intelligent Network approach, the need for a sophisticated local central office switch is reduced—the switch need only to know when to ask for guidance and to follow instructions.

Figure 4.7.
Conceptual architecture of the Intelligent Network



After Robrock (1991).

The approach effectively decouples basic switching from service provision and allows network providers to develop their own services, independent of the evolution of the switch hardware. New features and services can be provided from the centralized data bases (known in the industry as *Service Control Points*, or SCPs) and thus rapid deployment can be realized. The migration of the switching logic to the SCPs also promises to allow the network providers to work with multiple vendors in the design and development of new calling features and services.

Specific implementations of Intelligent Network principles are being designed by Bellcore and the individual telephone companies. However, the significance behind the Intelligent Network concept lies not in the specific features or efficiencies that may result from its early incarnations. The Intelligent Network approach to network design makes a conceptual distinction between the basic switching function and the service logic, or network control function. It is within the control function that new, innovative calling services may develop. The separation of the service logic from the network switching hardware can be likened to the separation of functions in a personal computer. The network switch can be compared to the computer hardware and the service logic is analogous to the many application programs that can run on the computer. The notion, not only that these functions are separable, but that the network is most efficient when they are separate, is a highly significant change in the telephone industry's philosophy and could have dramatic effects on the evolution of the public network.

Broadband ISDN will adopt Intelligent Network principles for signaling and control. The routing of ATM cells will require a separate control network to request and establish the virtual connections identified by the cell headers. This disaggregation of the control logic from the local switches will enable greater user choice in the utilization of network control resources. Whereas before the user had a pre-defined signaling relationship with the network (i.e. picking up the handset brought dial-tone from the local switch), in B-ISDN the signaling channels will be requested by a process known as "meta-signaling." Each user is assigned a permanent meta-signaling virtual channel that enables the user to request a particular type of signaling channel. The meta-signaling concept adds a degree of flexibility by giving the user the opportunity to select signaling and control services from different providers. For example, a simple signaling network may be set up to process conventional telephone calls while other signaling networks may handle more complex calls, such as multimedia conferences.

Because Broadband ISDN will involve the provision of many different types of services, whether separately or on an integrated basis, the construction of a "call" is inherently more complex than for a telephone network. Depending on the type of call, multiple connections may need to be established. A user dialing up a multimedia bulletin board will likely require separate virtual channels for the video, audio, image, and text components of the presentation. Similarly, two users conversing may wish, during the course of their conversation, to add different media elements to the discussion. For example, an advertising executive and a client might wish to view several candidate television commercials and discuss the merits of each. In this case, the users might add and drop their video connection several times during the discussion. In addition to the added complexity due to multimedia considerations, B-ISDN holds a potential for moving beyond the simple point-to-point, two-person call. The ease of communicating with multiple media will probably spur demand for conferencing applications, in which multiple users are periodically added and dropped from a call.[43] Furthermore, information sources will be distributed across the network, and thus it is easy to imagine that, in the example of the advertising executive, the video sources are not at one of the user's terminals, but must be called up from another location in the network. For both parties to view the video images, some form of bridging connection will have to be established.

The high degree of complexity in the B-ISDN call model renders current signaling techniques obsolete. Protocols developed for telephony and simple data applications are

---

[43] Multiple-party calls, in which identical information is distributed to multiple users, are facilitated by the "multicast" feature of B-ISDN. Multicasting involves the duplication of the payload of incoming cells and the routing of those duplicates to the intended destinations.

insufficient to handle some of the sophisticated signaling requirements associated with multimedia, multi-party communications. In the CCITT, signaling standards are being developed along two parallel tracks (Ransom and Spears, 1992).[44] The first effort involves an attempt to upgrade the capabilities of the Q.931 user-network signaling protocol developed for N-ISDN. The second effort is a longer-term, fundamental rethinking of the call model and the development of a new signaling approach for B-ISDN. The two-track effort implies a phased introduction of signaling for B-ISDN networks. Initially, signaling may be quite simple and limited in capabilities. Indeed, given the relative sluggishness of the CCITT standards process, it is entirely possible that ATM services may debut in the public network without any signaling capabilities. If premise-based ATM switches develop signaling protocols independent of the CCITT standardization efforts, then demand may blossom for public ATM switching of semipermanent virtual connections. In this scenario, users connect premise-based ATM switches with virtual paths and signaling is performed outside of the public network. This phase could be followed by the introduction of limited public signaling services and, eventually, sophisticated signaling capabilities would become available.

4.4.5 Traffic and Congestion

Traffic engineering is far more simple when all channels are based on a single fixed rate, as is the case with a telephone network, than in an ATM environment, where the channel capacities range by orders of magnitudes and the data streams can come in varying degrees of burstiness. The process of requesting service in a telephone network is also relatively simple—one lifts the receiver and, if a dial tone is returned, can proceed to enter a telephone number

---

[44] In addition to CCITT efforts, communications equipment manufacturers are developing industry standards for private ATM networks. For example, the Fiber Distributed Data Interface (FDDI) Follow-on LAN project (Fink and Ross, 1992) is investigating the interconnection of FDDI-based local area networks with ATM switching. The ATM Forum, whose member companies include the equipment manufacturer Northern Telecom and the carrier US Sprint, is developing standards for the interconnection of computer workstations using ATM. If the CCITT is unable to arrive at signaling standards in timely fashion, these industry groups may be forced to develop their own signaling standards.

designating the intended party. With the unpredictable nature of broadband traffic requirements, the CCITT has accordingly recommended a rather more complex approach to network congestion that involves a "negotiation" between the user and the network.

Unlike circuit-switching, where, for the duration of a call, a dedicated communications path is established between the end users, the resources in an ATM network are allocated on a virtual basis. In establishing a virtual connection, the network must have a reasonably accurate idea of the amount of traffic (i.e. number of 53-byte cells per unit time) to expect from the user requesting the connection. Based on this knowledge and knowledge of the traffic being generated by other users in the network, the network can determine whether it has the resources to meet the user's request. Note that in a packet-switched network, traffic congestion does not result in delayed dial tone, "all circuits are busy" warnings, or busy signals; instead cells are either delayed en route or simply lost along the way,[45] leading to quality degradations that can be harmless or severe, depending on the application. In order to minimize the potential for such degradations, the CCITT has developed the concept of service negotiation. In this arrangement, the user, when requesting a virtual connection from the network, must not only specify the destination but must include information on the source traffic characteristics and the required quality of service class. Source traffic may be characterized by a number of different parameters—standardization is not complete in this area—such as average bit rate, peak bit rate, burstiness, and peak duration. The quality of service (QOS) indication is still a subject for further study within the CCITT. It is likely that network providers will offer a menu of QOS classes, defined by such parameters as cell loss rate and variation in cell delay, from which the user may

---

[45] The effect of cell loss can be mitigated by the use of the Cell Loss Priority (CLP) feature of B-ISDN. Within each virtual channel, users may label some cells with priority status. In the event of heavy congestion, cells without this priority marking would be discarded first. This feature is particularly useful in applications, such as layered video coding (Verbiest *et al*, 1988), where some cells contain more important information than others. Note that the CLP feature refers to cell priority within a virtual channel only—overall network priority can be established at the virtual channel or virtual path level (i.e. cells associated certain some virtual channels could have priority over those linked to other virtual channels).

choose. Within the network, the Connection Acceptance Control (CAC) function processes connection requests and allocates the necessary resources, if available for the connection.

In the discussion of service negotiation, it is axiomatic that the individual users cannot be expected to perform the actual negotiation with the network CAC processor. The negotiation will undoubtedly be mediated by the user terminals or intelligent processors from the network provider or from other service providers. The service mediation concept is analogous to front-end, user-friendly interfaces in personal computers. Just as Microsoft Windows serves as a front-end for the MS-DOS operating system on many personal computers, translating the user's wishes into syntactically correct DOS commands, user interfaces within intelligent multipurpose workstations, videoconferencing codecs will mediate the service negotiation between the user and the network. Alternatively, such mediation services could be provided through intelligent processors operated by independent service providers, or by the network carriers.

The outcome of the service negotiation can be viewed as a contract between the user and the network. This contract, like any other, must be enforceable. To ensure that once the parameters of a connection have been negotiated, they are not violated, the CCITT has developed the rather sinister-sounding concept of the network policing function. The policing function simply monitors the traffic being generated by the user to verify that it is not exceeding any of the parameters that had been promised in the negotiation. The particulars of how violations shall be defined have not yet been standardized by the CCITT. For example, a suitable period over which to measure the average bit rate must be established. Suitable responses to violations are also under study. Alternative police actions include simply dropping cells that exceed promised rates or dropping the entire connection.

Chapter Five

ATM Networks I: Service Independence

5.1 INTRODUCTION

The most significant implication of the CCITT's choice of ATM as the basis of Broadband ISDN is that, by the time the public telephone network has evolved into a public broadband communications network, it will no longer be recognizable as a telephone network—it will be something entirely new. In order to understand the nature of future public broadband networks and how they may develop, one must draw a distinction between Broadband ISDN, on the one hand, and ATM, on the other.

Broadband ISDN is a vision. It is what the world's telephone companies, under the auspices of the CCITT, saw as the ultimate extension of the telephone network, made possible by the technological advances in high-speed electronics and optical fiber. In this vision, all communication services would be carried by one universal super network. The vision is bold, even audacious, in that it seeks to develop a single network that is capable not only of supporting all extant forms of telecommunications, but even all of those envisaged for the future. The breadth of services identified by the CCITT for Broadband ISDN provides the first clue that this new network will have transcended mere telephony.

Asynchronous Transfer Mode is the network technology, or, more specifically, the switching and multiplexing technique chosen to form the basis of this Holy Grail of networks. It is the vehicle that enables the cornucopia of services envisaged for B-ISDN to travel in harmony across a single network infrastructure. While ATM is frequently linked to B-ISDN, it is important to recognize that ATM is a method of transferring information that is independent of B-ISDN. Although the CCITT, by choosing ATM for its B-ISDN vision, may have been an early promoter of ATM, it is the data communications equipment manufacturers who are developing the first wave of ATM products. Indeed, it is likely that the first implementations of ATM switching will be found in customer premises networks, not in the public network. Similarly, one may find that early public network applications of ATM will not involve the provision of new services to end users so much as to offer new types of connections between nodes in a network.

In order to understand how public broadband networks will develop, it is important to focus not so much on the vision, which is B-ISDN, as on the means, which is ATM. The vision may or may not be realized. Deployment of ATM technology, however, is assured—ATM switches for private networks have already been announced.[46] ATM was chosen as the solution for the ambitious requirements of B-ISDN because of its flexibility—its ability to support communications of varying characteristics. The key to ATM's flexibility lies in its fundamental indifference to the applications it supports. At one point in its B-ISDN deliberations, the CCITT investigated the possibility of using multirate circuit switching, a technique which supports a finite number of different fixed rate channels (De Prycker, 1991). Bit rates would be set for a variety of speeds, ranging from low, such as the 64 Kbps used for digital transmission of voice signals, to high-speed channels, perhaps on the order of 40 Mbps for video signals, and each service would have to be adapted to the channel that fit the best. ATM is based on a dramatically different approach—rather than building a network that is capable of supporting a variety of specific applications, such as voice, high-speed data, and video, the CCITT chose to develop a network that is flexible enough to support all applications.

---

[46] See "Adaptive Unveils ATM Switch, Ready in Fall," *Communications Week*, July 13, 1992, p.4, col. 1.

The flexibility of ATM is based on the segmentation of digital data from any communications application into 53-byte cells. The ATM cell is universal—every cell, whether it is used to carry a snippet of a voice, a picture, a financial transaction, or a symphony, not only looks the same, but is given the same treatment by the network. This cell-based approach to information transport is the embodiment of the saying that "bits are bits," regardless of what they signify. The transport and transfer of information across an ATM network is in this sense truly service-independent.

The concept of service independence must be contrasted with the vision of service integration. An ATM network is best understood as a communications platform, one that enables the provision of multiple services, rather than a network that actually provides multiple services on an integrated basis. The latter type of network would be one in which the network had multiple interfaces to support different applications. The network would be capable of adapting to each different application, such as voice or video, and transporting it from end-to-end. In this integrated service network, the nature of each application is relevant to its treatment by the network—voice and video have different transmission requirements and thus demand different processing. In a communications platform network, the interfaces to the network are service-independent and the task of adapting the specific application to the network interface rests with the terminal equipment. The network itself, because all applications conform to its interfaces, does not concern itself with the application. ATM, because it mandates the segmentation of all user data into 53-byte cells, and is therefore indifferent to the applications it supports, is such a communications platform.

The skeptical reader will note that the CCITT, despite having developed a service-independent network, is nevertheless standardizing the ATM Adaptation Layer (AAL) functions that will accommodate different applications. This standardization effort, while organized by the CCITT, does not change the fact that the ATM Adaptation Layer functions will be performed at the edges of the network and not within the network itself. The interfaces to the network will be defined through the ATM layer and thus, while the CCITT standards for AAL functions may serve as useful guides for adapting different applications to the ATM network, users are not bound to CCITT standards—they may develop their own AAL techniques. This is not to say that standards for AAL functions will not be necessary—if communications devices from different manufacturers are to interoperate across an ATM network, they will need to follow common procedures for adapting user information into ATM cells. The relationship between the ATM layer and the ATM Adaptation layer functions is analogous to the PSTN and the fax machine. The public network interface is the telephone jack and the coding of an image into the signals that can be transported across the PSTN is a form of adaptation. While standards for the coding and transmission of facsimile images need not have come from the network providers—it could have come from the electronic equipment industry, for example—the need for standards was nevertheless critical to the emergence of the market for facsimile applications.

5.2 NETWORK APPLICATION SCENARIOS

The service independence of an ATM-based platform network is illustrated in the following application scenarios. Three general applications have been chosen for demonstration: telephony, high-speed connectionless data communication, and video program retrieval. The examples are highly simplified and do not necessarily represent preferred solutions. They are offered as examples of how each of the applications could be supported by an ATM-based platform network.

5.2.1 Application I: Telephony

A stylized local exchange network, circa 1992, is shown in Figure 5.1. Telephone trunks connect three central offices. These trunks consist of 64 Kbps channels bundled into T1 circuits, which are further bundled into 45 Mbps DS-3 circuits, and possibly carried in SONET frames. The transmission medium supporting these circuits is optical fiber. Each office is linked, via separate low-speed data circuits to an Intelligent Network node for call set-up and other call control operations.[47] Most subscribers connect to the central offices over dedicated copper pairs, though some subscriber lines, as shown at the bottom of Figure 5.1, are multiplexed over T1 or DS-3 facilities.

Figure 5.1.
A simplified depiction of a contemporary local exchange telephone network.

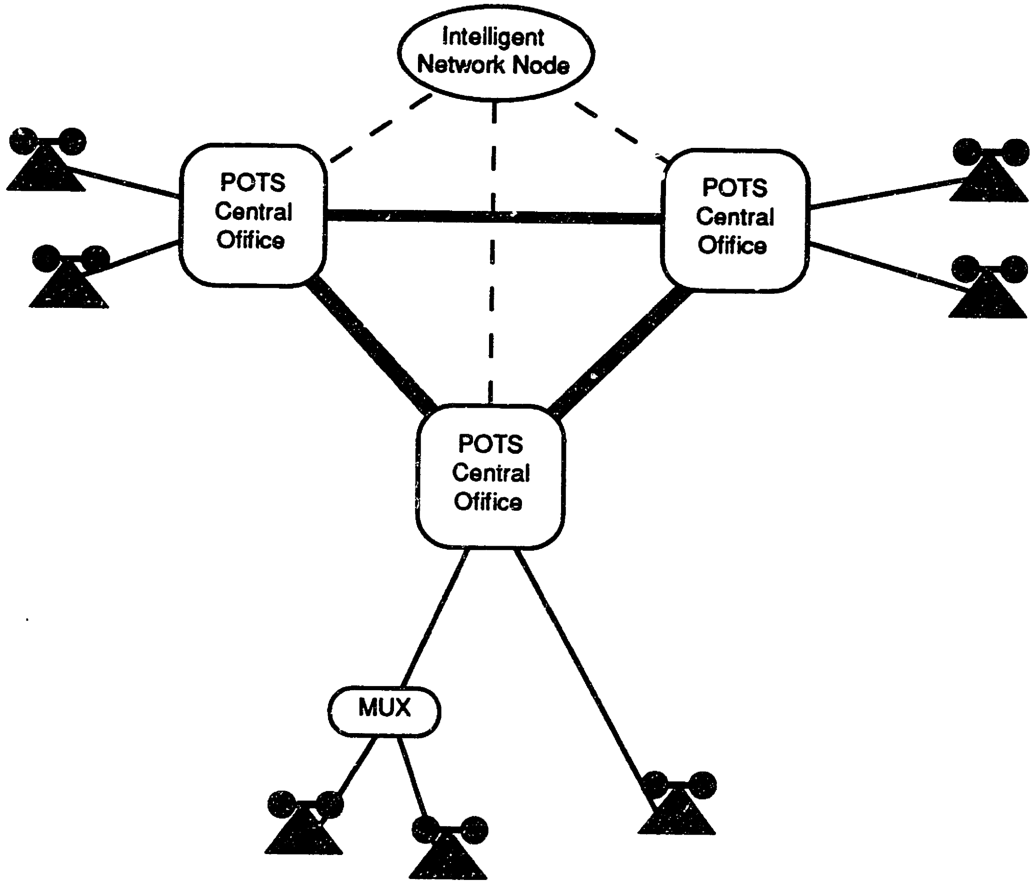


Figure 5.2 shows an early phase of the introduction of ATM into the network. The most notable difference from the contemporary network is the replacement of the direct trunks

---

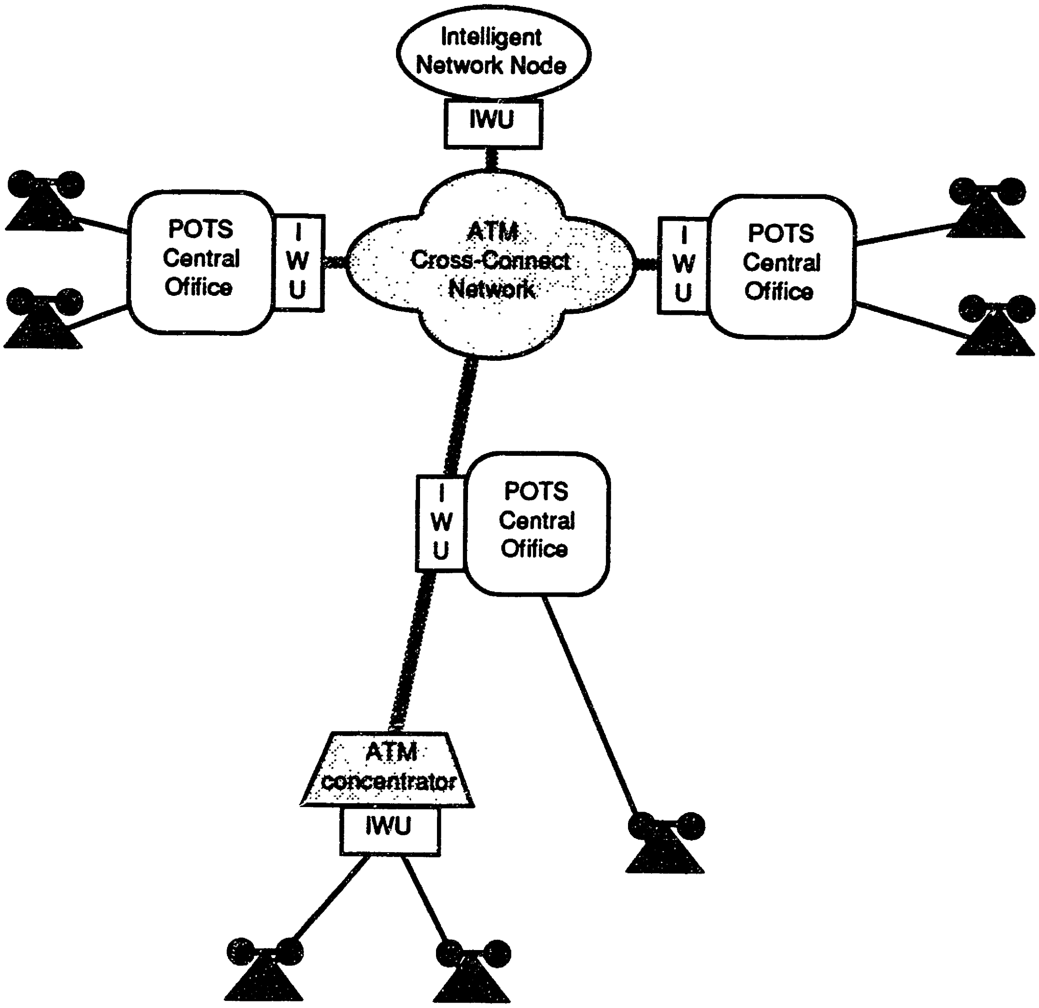
[47] The use of a single node for the Intelligent Network is a gross simplification. In reality, this "node" is more likely to be a network of intelligent processors.

between each central office. In this new network configuration, an ATM virtual path cross-connect network is used to provide *virtual* trunks between each office. The same central offices shown in Figure 5.1 continue to be used—they most certainly will not be replaced overnight by ATM switches—and have been augmented with ATM inter-working units (IWUs). The role of the inter-working unit will be significant in early ATM applications. An inter-working unit performs the ATM Adaptation layer functions and enables existing equipment and services to interoperate with ATM transmission. Inter-working units are not universal—they will be designed for specific adaptations between ATM and non-ATM elements of the network. In this example, the inter-working unit performs the conversion of constant streams of digital voice signals into ATM cells and provides for the end-to-end timing necessary for smooth voice communication.

In the subscriber network, the multiplexer has been replaced by an ATM concentrator. Unlike the standard T1 or DS-3 multiplexer, the ATM concentrator does not provide a dedicated path back to the central office for each voice circuit, but performs a statistical multiplexing function, aggregating the ATM cells of different subscribers onto a common path. The links to the Intelligent Network have also become more efficient—dedicated data circuits are replaced by semipermanent virtual channels connecting each central office switch to the control network.

Note that in the configuration shown in Figure 5.2, one can draw a distinction between the ATM network provider and the telephone service provider. The boundary of these two networks lies at the interface between the ATM network and the interworking-units. From this perspective, the telephone service provider can be viewed as a value-added provider, utilizing some of the basic transport capabilities, namely virtual path and virtual channel connections, of the ATM network to support the interconnection of the nodes in its telephone network. One can draw an historical analogy to the telephone companies and the value-added networks, the providers of packet switching services. The VAN providers used telephone company facilities, both switched and dedicated lines, to connect their packet switches and to enable users to reach their access nodes. In this case, the telephone service provider uses basic network facilities to connect its telephone switches.

Figure 5.2.
Early introduction of ATM into the telephone network.



ATM virtual paths are used to connect the central offices and an ATM concentrator is used to transport some subscriber traffic to a central office.

A long-term scenario is shown in Figure 5.3. In this configuration, all POTS telephone switches have been replaced by ATM virtual channel/virtual path switches. The Intelligent Network concept has been implemented to its logical extension and telephone service is provided, not so much by the ATM switches, but through separate control networks. Users request calls by interacting with telephone service control networks, which, in turn, request the establishment of appropriate virtual channels from the network provider's connection control network. The latter control network services requests for virtual channels and is indifferent to their use—it has no knowledge that the channels requested are to support telephony.

Multiple service control networks are shown in Figure 5.3. These control networks, whether they support telephone service or multimedia information services, have similar relationships to the ATM network provider. Each requests certain types of virtual connections from the ATM network on behalf of end users and does not need to own or operate the ATM switching or transmission facilities. Telephone service thus becomes what we now consider an "enhanced" service—it is a higher level service that operates over a "basic" network platform. The traditional *Computer II* distinction between basic and enhanced services has, in this network topology, completely disappeared.

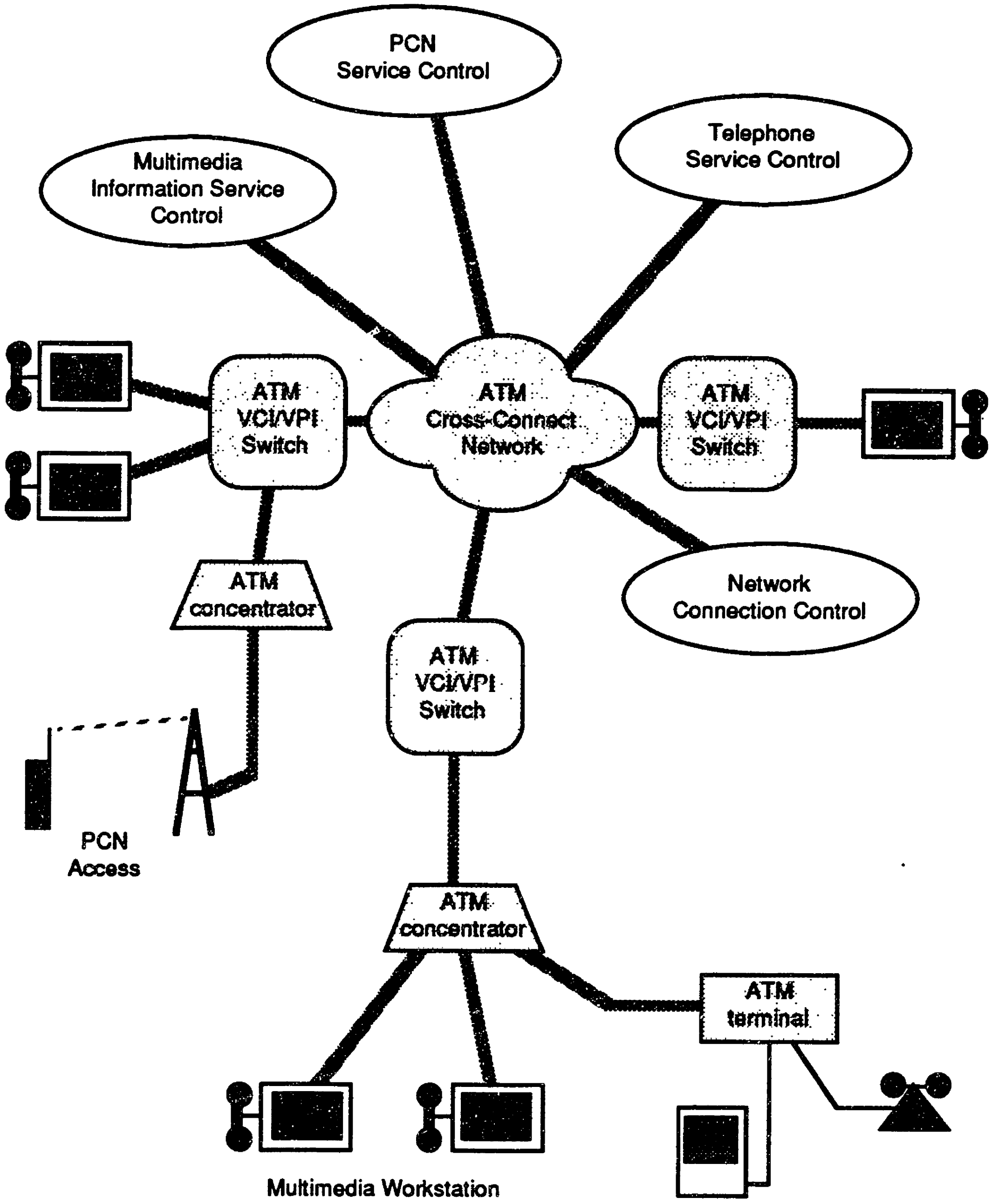
Figure 5.3 also depicts the use of an ATM platform network to support transmission of telephone traffic from a mobile personal communications network (PCN).[48] Like conventional, wireline networks, PCNs will also need to transport trunk traffic between different radio access nodes, between central office switches, and to and from the long distance network. The PCN service is shown to be offered by a competing carrier, with a separate control network. The example shows how competing telephone service providers can distinguish themselves on the

---

[48] PCN, for Personal Communications Network, and PCS (Personal Communications Services) are terms often used interchangeably to describe wireless communications networks that will support telephony and other applications. For general discussions of PCN/PCS, see Bryan (1991) or Singer and Irwin (1991). For a more technical treatment of wireless communications trends, see Goodman (1991). PCN/PCS technology and policy issues are covered extensively in the February, 1991 and June, 1992 issues of IEEE Communications Magazine.

basis of their access methods (wireless or wireline) and the software-based calling features in their control networks, while utilizing common elements of the basic ATM platform network.

Figure 5.3.
The long-term scenario: telephony and other applications provided over a total ATM network.



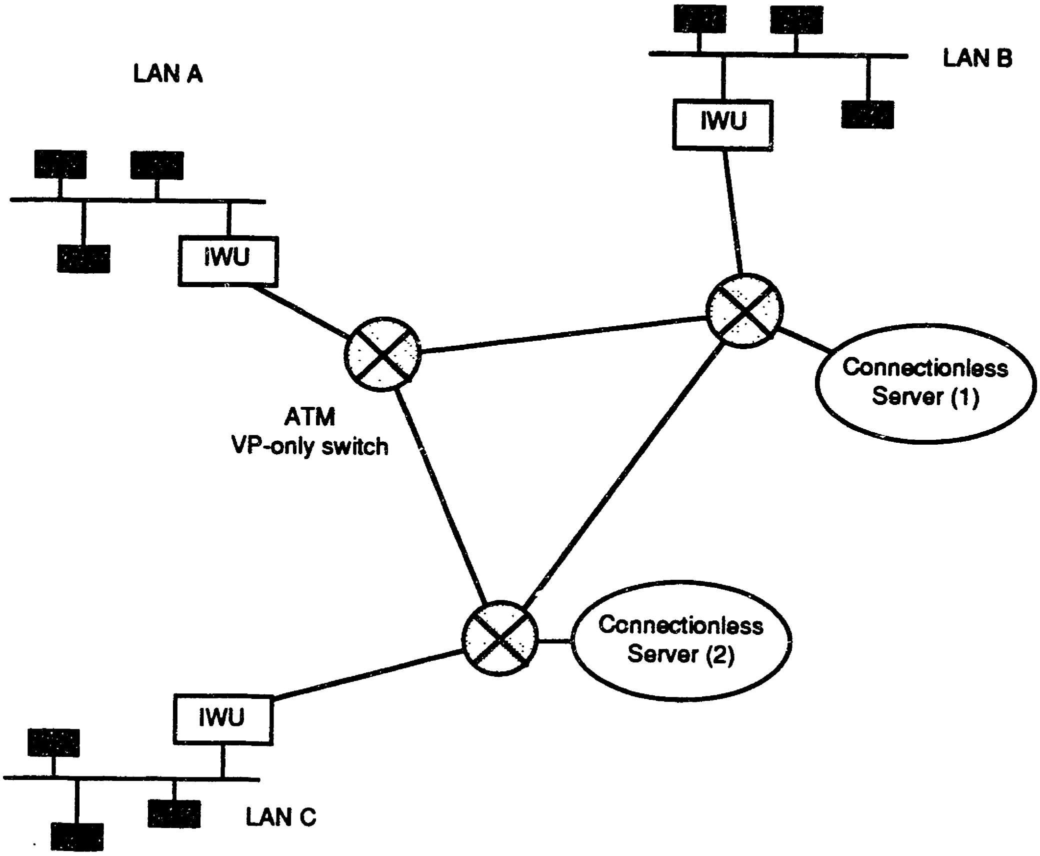
5.2.2 Application II: Connectionless Data

ATM, as noted in Chapter 4, is a connection-oriented service. However, many data applications, especially high-speed, bursty applications, are most efficiently served by connectionless routing techniques. In a connectionless service, "calls" need not be established, for each burst of data contains explicit routing information. While ATM is a connection-oriented service, connectionless services can be provided over an ATM network as shown in Figure 5.4 (Lobjinski *et al*, 1990). In this example, two connectionless servers, the processors that route incoming data packets to appropriate network addresses, are used to support the interconnection of users on geographically separate local area networks. The connectionless servers, like the telephone switches in the previous example, are not so much a part of the basic network as they are processors located on the periphery of the network.

Each local area network is connected via semipermanent virtual paths to a connectionless server. The two connectionless servers are connected in the same mode. Inter-working units are needed at each LAN and each server to adapt the connectionless data packets into ATM cells for transport across the network. In the configuration of Figure 5.4, LANs A and C are served by server 2 and LAN B by server 1. Data packets destined from LAN A to LAN C travel over a virtual path from LAN A to server 2, where they are routed along a virtual path to LAN C. In the case of traffic between LAN A and LAN B, packets travel along the virtual path to server 2, are routed along a second virtual path to server 1, and finally along a third path to LAN B.

The example of connectionless data service highlights the use of an ATM network as a platform. While the high-speed transfer of data can be performed by pure ATM switches, users are not limited to ATM techniques—those desiring the particular advantages of connectionless services or other transfer techniques can use the ATM virtual path network to reach optional switching resources.

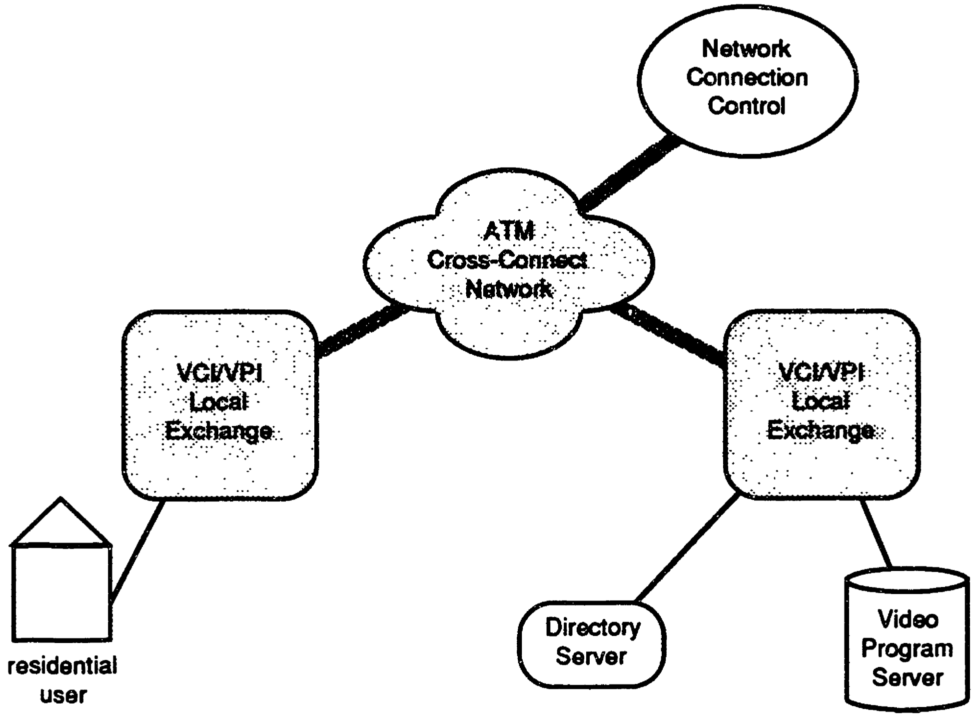
Figure 5.4.
Interconnection of local area networks using a connectionless data service provided over an ATM virtual path platform network.



5.2.3 Application III: Video Program Retrieval

A video-on-demand application is shown in Figure 5.5. A residential or business user with an ATM-compatible, multimedia terminal searches the directory of a video programming service provider (analogous to a video rental store) and selects a video. Once the user expresses his or her choice, the video program is presented for viewing.

Figure 5.5.
Provision of residential video on demand services using a switched ATM network.



The user begins the session by establishing a data connection to the directory server of the service provider. After browsing through the directory, the user selects a particular program. This choice leads to a request to set-up a channel between the user and the video program server. Once the channel is established, the user may begin to view the program, and, with a separate data connection also established with the video server, control the presentation of the program.

This example, which depicts the function of call by call switching in an ATM network, also serves to underscore the importance of signaling. As in the telephone network, some form of signaling is needed to enable the user to communicate the desired destination and for the network to establish an appropriate connection. Furthermore, in the video retrieval case, the process of searching for a program, requesting its delivery, viewing the program, and controlling the presentation of the program is quite complex, requiring several virtual channel connections. If such a service is to succeed, the user will have to be shielded from this complexity. One may therefore see implementations of this application where the connections between the user and the video server are requested, not by the user, but by the directory server, in a brokered manner.

As is the case in the two previous examples, the ATM network in this scenario is unaware of the services (i.e. data communication, synchronous video) it is transporting. Even though signaling is involved in this application, the service is irrelevant—the negotiation between the user and the network signaling entity for the virtual connection makes reference not to video and data, but to parameters such as bit rate and quality of service.

5.3 CONCLUSIONS

The above three examples of communications applications supported by an ATM platform network illustrate the service independence of ATM. Not only does the network appear indifferent to the applications it supports, it is noteworthy that the basic capabilities of the network are used similarly by the different applications. In effect, the ATM network provider is a wholesaler of basic commodity products that serve as inputs to the development of the products of the service providers. A consequence of this view is a new conception of a telephone network as one of many specialized networks that utilizes a more fundamental platform of ATM network services. The telephone network is no longer a basic communications network that supports other services, it is in fact one of the other services.

The telephony scenario discussed above distinguishes the ATM network provider from the telephone service provider. This distinction is not prescriptive—it does not imply that the network provider and the telephone company must be separate corporate entities nor does it advocate a structural separation. The distinction is functional in nature, reflecting both a conceptual separation and a technologically feasible separability of the functions performed in advanced telecommunications networks. A more detailed analysis of the principal functions of a telecommunications network follows in Chapter 6.

Chapter Six

ATM Networks II: Network Functions

6.1 INTRODUCTION

In the past, the functions of the public switched telephone network could be simply divided into switching and transmission (Rey, 1984). Switching enabled one user to connect to another without requiring a dedicated circuit, while transmission was used to connect the users to the switches and, in turn, the switches to other switches. Signaling, which was needed to enable the users to communicate with the switches and the switches to communicate with each other, was embedded in the switches. In the early days, call processing, or the treatment of call requests, was limited to connecting the caller to the number called. As described in Chapter 4, the advent of stored program control switches sparked the development of software features to provide enhanced call processing. Custom calling features proliferated and the call processing duties of the central office switch accordingly became more complex. Switching was thus a very broad function, encompassing both the physical connection of circuits and the call processing intelligence needed to make the proper connections. With the adoption of Intelligent Network principles, however, these two switching functions have been separated. The control logic has been removed from the local switches, both physically and functionally.

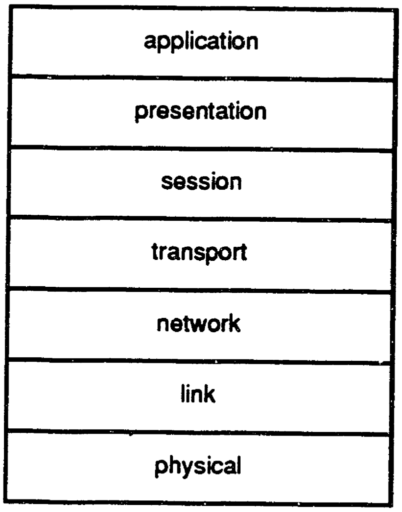
The ATM approach to switching takes this disaggregation of functions a step further. Practically all that is required of an ATM "switch" is to read the header bytes of an incoming cell, look in a routing table to determine the appropriate output port, and pass on the cell. Periodically, the control network swoops down and changes the values in the routing table. The simplicity of the ATM switch is actually critical to the ability of ATM to achieve the high processing speeds envisioned for B-ISDN. The cell routing of an ATM switch can be implemented entirely in hardware—no software processing is needed. The intelligence of the system rests in the control network that must devise the end-to-end virtual connections across a network of ATM switches. The disaggregation of switching functions is not limited to the separation between the control network and the ATM switch. The concept of virtual paths and virtual channels allows for two levels of ATM cell routing. Some network components can serve as virtual path-only switches, while others may route cells at both the virtual path and virtual channel levels. It is important to recognize that this disaggregation of functions is not merely for internal efficiency's sake, but the subdivided functions represent, in some cases, building blocks that can be combined or used individually to provide services.

The effect of this disaggregation of switching functions into multiple components is to distort the meaning of the term "switching." One can no longer speak of switching monolithically, as in a central office "switch." Instead, one finds several different switching functions, that, when taken as a whole, provide an integrated service similar to the switching provided in the plain old telephone network.

6.2 NETWORK FUNCTIONS

In understanding the functions of a communications network, it is often instructive to think in terms of layers. Each layer of communication performs functions that enable communications to take place at higher layers. The International Standards Organization's Open Systems Interconnection (OSI) reference model (Figure 6.1) is the archetypal model for layered communications. The delineation of the different layers in the OSI model is dependent on defining the functions that take place within each layer. While the OSI model provides a constructive perspective from which to view communications, the layers it presents are not universally applicable—some communications protocols are based on layers that do not correspond directly to those in OSI. In analyzing the component functions of an ATM-based network, one can describe several functional layers, supplemented by functions that interact at multiple layers.

Figure 6.1.
The OSI Reference Model



Source: Jain and Agrawala (1990)

The lowest, most fundamental, layer of a modern communications network is the physical medium itself—the optical fiber, the copper wire, or the radio spectrum along which communications signals travel.[49] The physical medium supports the transmission of signals

---

[49] One could, in the case of terrestrial media such as copper or fiber, extend the layered concept to include, below the medium, the duct that carries the cable, or even the right-of-way that allows for the duct to be buried. When discussing policy matters such as access to communications networks at different levels, it may in certain cases be useful to include these additional layers.

between the nodes of a network. The higher layer functions of an ATM-based network are outlined below.

6.2.1 Transmission

Conceptually located directly above the physical medium, transmission is the carriage of communications signals along the medium. In today's increasingly digital world and most certainly in an ATM-based network, transmission is the conduit of bits between nodes of the network. For years transmission capacity has been organized into bundles of 64 Kbps voice channels. With the introduction of SONET, the transmission network can accommodate a greater variety of data rates, but is still oriented towards the provision of fixed rate channels. A key feature of transmission is the provision of what are known as "clear" channels, or "bit pipes" that enable unadulterated, end-to-end communication. What comes out of a clear channel is exactly what went into the channel. Transmission, it must be noted, is entirely independent of the ATM used to transfer cells between different network nodes—transmission capacity can be used to transport ATM cells or data streams organized in any other fashion. Like the ATM platform network that is indifferent to the service it supports, the transmission function in a network is indifferent to the organization of the data streams it carries.

The transmission function can be subdivided, not vertically, as in the layered model, but rather spatially, into *access* and *transport*. Access, in this model, is defined as the transmission between a user and the network, whereas transport refers to transmission within the network. The connection between the user and the network has for many years taken the form of a pair of copper wires that run from a subscriber's premise to the network provider's central office switch. However, in recent years, with the decentralization of central office switching, in the form of remote switching units, and the introduction of active multiplexing equipment in the subscriber loop network, the boundary of the network has shifted and the definition of access must be updated accordingly. A less technology-dependent conception of access would be to relate it to the portion of the network that serves a single customer. Thus the portion of the network between the user and the first SONET multiplexer, or between the user and the first ATM concentrator, would be considered the access portion of the network.

Transport, quite simply, is all transmission that is not access. It is the provision of communications channels between nodes of the network. In this definition, the use of the term network is not presumed to be monolithic. The "network" is, of course, a patchwork of many interconnected and independently operated networks. For example, in a contemporary telephone network, transport would refer to both the transmission circuits between two central offices of the same local exchange carrier and the circuits between a local exchange carrier's office and the point-of-presence (POP) of an interexchange carrier. Note also that, by this definition, transport would include the transmission of cells between an ATM concentrator and an ATM VCI/VPI switch and between two ATM VP-only switches.

The distinction between access and transport is not merely logical; the two functions are physically separable. For example, Figure 5.3 showed a PCN form of access, where a radio transmitter served many individual telephone handsets. The radio transmitter serves as a well-defined boundary between the access and the transport functions of the network—using a standard interface at this boundary would enable the two functions to be performed by separate companies. Furthermore, insofar as the radio equipment can be separated from the network transmission equipment, the costs associated with each function are also separable. Another example of the boundary between access and transport is the interface at a Fiber-to-the-Curb[50] pedestal that converts optical signals associated with multiple subscribers into individual copper or coaxial cable drops.

---

[50] For an overview of Fiber-to-the-Curb systems, see Shumate and Snelling (1991).

6.2.2 Transfer

As discussed above, the term switching has lost much of its precision and must be understood in terms of its component functions. One of these basic functions can be called *transfer*. The CCITT adopted the term transfer, as in Asynchronous *Transfer* Mode, in recognition of the fact that the simple relaying of cells across a network fell somewhere between conventional notions of switching and transmission. In the layered model, transfer is one level above transmission. The transmission network provides the dedicated clear channels in which ATM cells travel. At ATM transfer nodes, cells traveling within incoming transmission paths are examined and actively directed towards appropriate outgoing transmission paths.

Virtual Paths

The transfer layer is subdivided, per CCITT standards, into the virtual path sublayer and the virtual channel sublayer. The virtual path sublayer, which typically carries a bundle of virtual channels (Figure 4.2), is the lower of the two sublayers. Virtual paths are typically envisioned to be configured on a semipermanent basis. Rather than creating virtual paths on a call-by-call basis, users and network providers will request the establishment of virtual paths between different endpoints of the network. These paths will remain available on a subscription basis and, as with individual call establishment, be defined to support certain traffic characteristics. While the virtual paths do involve the active routing of communications on a cell by cell basis, the functional difference between virtual path provision and the establishment of dedicated transmission channels is subtle. One could almost view the decision to establish a transmission channel on a dedicated or a virtual basis as an internal network decision. If a virtual path is configured to support fixed bit rate transmission and the network guarantees the speed and the clarity of the virtual path, does the user see any difference between this service and a dedicated clear channel service? In practice, probably not. Nevertheless, the logical separation is warranted. A virtual path network is necessarily provided on top of a transmission network, that is to say, the ATM cell headers containing the virtual path routing information travel within the dedicated channels provided in the transmission network. While at one limit, that of the virtual path configured to emulate a fixed rate clear-channel, the functions provided by dedicated transmission paths and virtual paths are indistinguishable, the inherent flexibility of a virtual path connection differs sharply from a transmission channel. Whereas a transmission channel is typically provided on a fixed rate basis, the virtual path can be configured for a range of data rates and for variable traffic patterns.

Virtual Channels

Virtual channels are distinguished from virtual paths in that they logically lie within virtual paths and because they are more likely to be created and destroyed on a call-by-call basis. For example, two ATM switches in a private network may be connected by a semipermanent virtual path; a call between two users associated with those switches would take place through virtual channels established within the virtual path. Virtual channel transfer thus lies on a higher level than virtual path transfer. As shown in some of the application scenarios presented in Chapter 5, VP-only switches can be used to connect users and network resources on a semipermanent basis and individual connections within these virtual path networks are routed on a virtual channel basis. The routing of cells according to their virtual channels must not be confused with switching—virtual channel transfer presupposes signaling transactions that have established the virtual channel connections across the network. Thus the function of a virtual channel/virtual path ATM "switch" remains quite limited since these switches merely route cells based on instructions they receive from other elements of the network.

Consistent with the relationship between virtual paths and transmission channels, virtual channel transfer can only be accomplished on top of virtual path connections, whereas virtual path transfer, at a lower layer, does not depend on the establishment of virtual channel services.

6.2.3 Control

The Intelligent Network concept removes some of the control of the switching process out of the central office switch and places it in remote network resources. The CCITT's development of the ATM concept takes this intellectual separation of transfer and control to an extreme, removing the control function to such an extent that it almost seems to have been abstracted away. In fact, the mysterious control network that is responsible for determining the routes for end-to-end virtual connections in an ATM network and communicating this routing information to the individual ATM cross-connects and ATM switches has been slow to emerge from the CCITT's deliberations on broadband signaling. As noted in Chapter 4, signaling and control in a broadband, multimedia network is as complex as the routing of ATM cells is simple. Furthermore, this complexity must be contained within the network—if it spreads to the human interface, then the network becomes virtually unusable. In order to manage the complexity of the control network, it has been conceptually divided into two separate functions, or functional modules, known as *connection* control and *call* control (Beau *et al*, 1990).

Connection Control

Connection control refers to the establishment, maintenance, and release of virtual connections in an ATM network on a call by call basis. The connection control function essentially processes each connection request independently, knowing nothing of the possible relationships among individual connections. In the example of a multimedia call, separate virtual channels are typically established for the different service components of the call. The voice communications are supported by one channel,[51] video images by another, and the text and/or graphics associated with a visual scratch pad travel within a third. The connection control function is not responsible for the linking of these different connections into a single

---

[51] ATM connections are actually unidirectional, thus a two-way voice call would actually require two virtual connections. For the purposes of this particular illustration, the two channels can viewed as a whole.

communication, or conversation—it merely responds to each request for connection establishment.

Implicit in the responsibility for establishing and maintaining the virtual connections in the network is a duty to monitor and manage traffic congestion in the network. In order to maintain a virtual connection at the level of service promised in the service negotiation, the network must avoid traffic jams that can degrade performance. The approach to traffic congestion, as detailed in Chapter 4, is to regulate the flow of cells at the edge of the network, only admitting connections whose traffic characteristics can be supported by the network, given existing usage conditions. The connection control module performs both the connection acceptance control and policing functions described in Chapter 4.

Once the connection requirements have been identified through a signaling transaction, connection control must determine a suitable path through the network. Having mapped out this route, the connection control module must provide routing instructions to each node in the ATM network, associating the virtual channel and virtual path identifiers in the ATM cell headers with physical routes through the network. These connection identifiers must also, of course, be communicated to the user terminals or other entities requesting the connections for the purpose of labeling the ATM cells to be transported.

Insofar as the connection control module must maintain virtual connections, it is also responsible for assimilating information about node failures and other potential roadblocks in the network and re-routing the connection around these trouble spots.

Call Control

Connection control makes it possible for users who are not linked by a semipermanent connection to communicate. Call control is required to make the network usable in an everyday manner. Call control associates the virtual connections in an ATM network with calls being requested by users, tracking the relationships among the different users and the different service components of a call. To return to the example of a multimedia conversation, the call control function translates the request from the user or users initiating the conversation into connections it must request from the connection control module. In this sense, the call control module acts as a systems integrator—it is responsible for assuring that the users in the call are presented with a fully functional multimedia interface and can therefore converse easily.

The features that one usually associates with a telephone network are the province of the call control module. In a personal communications application, where a user is associated with a single telephone number, one that can reach the user at his or her desk, in the car, or on the street, call control contains the intelligence to "track down" the user, checking a database to find appropriate routing instructions.

One way to understand the call control function is to view it as an interpreter, enabling a service transaction between the network and its users. The language of the service transaction is technical, involving terms such as average bit rate and burstiness factors, whereas the desires of the user are expressed in more human terms. Although we have, as a culture, developed a tolerance for the use of technical codes such as telephone numbers, a more natural way to initiate a call may be to ask the network to connect you to the desired party. In other words, a desirable feature of the network would be to enable a caller to request, for example, "Please connect me with my brother Bob." In an ATM environment, the connection control function will not converse with the user in these terms, and thus the call control function is needed. In Chapter 4, the analogy of a personal computer was presented, drawing a parallel between the role of a user-friendly shell program and the mediation necessary to enable communication between a user and the ATM network. One can extend this analogy to include both the mediation and the call services functions, described in the preceding paragraph, of the call control module and compare call control to the universe of application programs that run on a standard personal computer. In this analogy, calling features and services correspond to individual application programs and the translation of network speak into human terms is analogous to the user interface of each program. Using a calling service would not be unlike using a computer program such as a word processor or a spreadsheet. The computer program acts on the user's wishes by interacting with the computer; the calling service would interact similarly with the network.

One of the most fundamental concepts of telecommunications—the call model—is being re-examined for multi-service, multimedia environments. In a conventional telephone system, a call is simple, using only one medium, voice, and requiring but a single, point-to-point connection. In an ATM network that allows the transfer of all types of information, one's conception of a "call" is limited only by one's imagination. The call control function in a network has a well-defined output at one interface: the link to the connection control module. Interoperability requirements will force standardization of this interface to a limited number of signals. The other boundaries of call control, on the other hand, are virtually unlimited. To be sure, some standard interfaces between call control modules and user terminals will be needed to enable interoperability, but the concept of call control is also bounded in one direction by the human interface. The potential for creativity and innovation in this interface between the network and its users is enormous.

6.2.4 Administration

Call control governs the dynamic configuration of the network, requesting the allocation of resources to support calls as they occur. *Administration* refers to the static configuration of the network, concerning the allocation of resources on a semipermanent or subscription basis. Network providers in contemporary networks use administration tools regularly, performing such tasks as updating the customer database in a central office switch or configuring a routing path in a digital channel cross-connect system. Administration functions act upon an ATM network at multiple levels. At the transmission level, dedicated channels can be configured through administration of the SONET multiplexers and cross-connects in the network. ATM administration involves the assignment of semipermanent virtual paths and, conceivably, virtual channels, through interaction with the connection control module. Updating information in call control databases is an additional administrative task.

While administration typically comes under the purview of the network provider, telephone companies have, since the mid-1980s, developed customer interfaces to some of their administration systems. These systems enable managers of large corporate centrex systems to update various elements, such as the dialing plan or the calling features, of the telephone company-provided centrex service. Such administration systems are, in effect, automated ordering systems—the customer simply bypasses the order process and the changes requested are implemented rapidly. This demand for user control over customer-specific network software is likely to continue for the services provided over an ATM network. Similarly, service providers creating networks that run over a public ATM network will no doubt want the capability to reconfigure their networks promptly and will thus demand access to the network provider's administration systems.

6.2.5 Maintenance

The maintenance functions of the ATM network take place at both the transmission and transfer levels. The basic responsibility of the maintenance module is to ensure the reliable operation of the ATM and transmission networks. Specific tasks include performance monitoring, defect and failure detection, system protection, failure or performance information, and fault localization. The maintenance functions monitor the network continuously to detect any degradation in performance, such as repeated bit errors, or mislabeled ATM cells. Any significant defects or failure in a particular node of the network may trigger the maintenance function to alert other nodes in the network and remove the defective node from service, thus attempting to preserve the smooth operation of the system as a whole.

The CCITT is developing Broadband ISDN standards that specify the particular maintenance tasks to be performed at each level. This standards development parallels a more general CCITT effort on producing a standard for management of telecommunications networks known as TMN, for Telecommunications Management Network.[52] TMN, still in its early stages of definition, will provide for standard interfaces for maintenance interactions at different network levels and for communication between different maintenance modules.

6.2.6 Billing

The billing function of a network lies mainly outside the network itself. Billing a user for services involves the processing of billing data into periodic bills to be distributed and the tracking of customer accounts. However, as distinct from other billing tasks, the generation of the billing data is linked to the actual operation of the network.

Charging for usage of an ATM network is a complex topic, to be treated in Chapter 7. The tasks that must be performed by the network to produce billing data are linked to the network provider's charging policies. For example, if users were to be charged for the network resources they request in the service negotiation, then the billing data could come directly from a record of the service agreement. However, if users were to charged for their actual usage, i.e. the number of cells generated during the connection, then a cell counter must be implemented and its data fed to the billing system.[53] Further complicating the issue is the concept of "punitive" pricing, wherein a user is "fined" for violating the parameters of the service agreement.

The billing function is properly viewed as a separate service function, with communication links to the connection control and administration modules. However, due to the uncertainty about ATM charging and the likelihood of multiple solutions, the nature of this interface remains unclear.

---

[52] General principles for TMN were set forth in CCITT Recommendation M.30, published in the 1988 Blue Book. CCITT Study Group IV is currently engaged in the further development of TMN principles and specifications.

[53] It has been proposed that the task of counting cells for billing purposes could be performed by the policing function, which will be monitoring traffic for the purpose of congestion control (Appleton, 1991).

6.2.7 Network Security

In the past several years, Americans have become increasingly sensitized to the vulnerability of the public network. Major network outages, such as the Hinsdale Fire of 1989, AT&T's Martin Luther King Day outage in 1990, the Signaling System 7 failures in June of 1991, and the AT&T outage in September of 1991 that crippled the air traffic control system in the New York area, have prompted businesses to develop disaster recovery plans and have led to Congressional hearings.[54] The 1991 SS7 failures, in which regional signaling networks collapsed under floods of erroneous maintenance communications, led to speculation that computer hackers had infiltrated the network.[55] The increasing dependence of the network on software and the vulnerability of software to viruses and other forms of infiltration has led to renewed concern about the reliability and the security of the network.

The description in this chapter of the network as a series of interconnected modules, largely under software control, suggests that the network may have a number of interfaces, or gateways, open to users and service providers alike. Some concerned with the security of the network caution that the opening of network software to enhanced service providers might seriously jeopardize the reliability of the network.[56] As more and more individuals are allowed access to network software, the potential for misuse of the network, through malice or incompetence, increases accordingly. If the network is to be opened further, as the discussion of separable network functions implies that it can, then the security function takes on a greater role.

The security function, in an ATM network, lies naturally at the interfaces between functions provided only by the network provider and those that are performed by multiple

---

[54] On October 1, 1991, the House Committee on Energy and Commerce, Subcommittee on Telecommunications and Finance held a hearing on the telephone outages. In addition, the Government Information, Justice, and Agriculture Subcommittee of the House Committee on Government Operations held hearings, entitled *Federal Communications Commission Efforts to Assure Telephone Network Reliability*, on July 10 and October 2, 1991.

[55] See "Telephone Technology Questioned After Failures," *The New York Times*, June 28, 1991, p. A16, col. 1.

[56] See, for example, National Research Council (1989).

service providers or users. Security, in this implementation, becomes a function of screening the requests that issue forth from external modules. For example, a service provider performing an administrative function, such as configuring its own virtual paths, should be prevented from configuring the paths of other users. A validation or verification scheme must therefore be embedded into any function that services requests involving network resources.

In this conception, network security is not a separable function along the lines of call control or network administration, but is rather woven throughout many of the functional modules of the network.

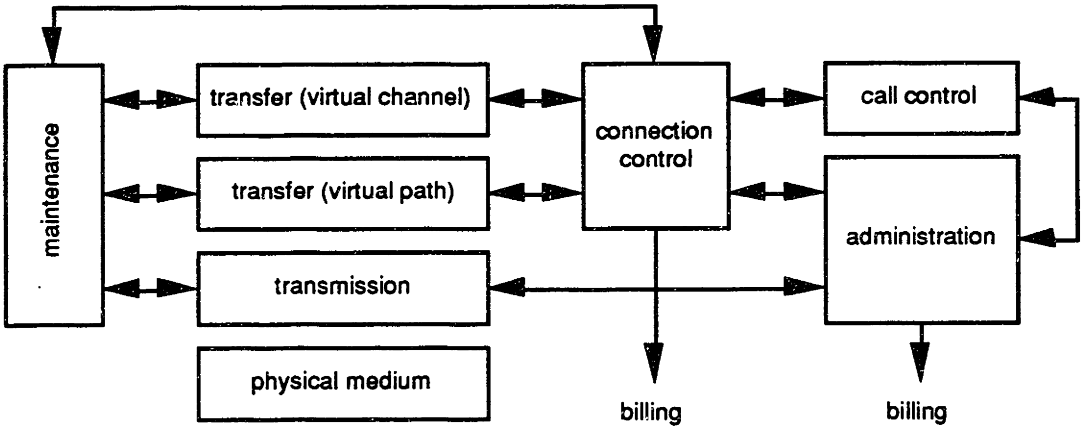
6.3 A CONCEPTUAL MODEL OF THE NETWORK

A conceptual model of an ATM-based network is shown in Figure 6.2. The functions described above are represented as separate modules connected by communications links. A key result of the ATM concept and the attendant development of a new network architecture and a new network philosophy for B-ISDN, has been the intellectual separation of the different functions that a network must perform. Old conceptions of the network resembled the paradigm of host computing—"dumb" terminals, i.e. telephones, were connected to massive host computers, the central office switches, that performed all of the processing duties. These processing functions were consolidated within each central office switch and, in general, the different switches did not share resources.

In this new conception of network design, a computer programmer's approach has been taken. A computer program typically contains a main body that delegates tasks to different subroutines. Each subroutine can be treated as the proverbial "black box" with a well-defined interface—the main body passes a set of parameters to the subroutine, and when the subroutine completes its processing task, returns new values for these parameters. The ATM-based network has been intellectually organized along the same paradigm. Network functions have been conceived as independent modules, each with specific tasks and with interfaces to other modules.

Note that this conceptual organization of the network is not linked to any specific equipment or physical locations. Call control modules could be bundled with connection control into an ATM virtual channel switch, or they could be distributed across the network as peripheral processors. Virtual channel routing and virtual path routing can be combined in a single device or virtual path routing can be performed independently. While the separation of functions is, for the moment, conceptual, it must be emphasized that it is also realizable. Efficiency considerations may prompt some manufacturers to build devices that combine many of these functions. However, as long as the interfaces between modules are standard and open, these modules are separable and their functions can be provided independently.

Figure 6.2.
A conceptual model of the network.



In analyzing the functions of the network, it is important to recognize that the functions, or modules, presented in this chapter, can be further subdivided. In particular, the peripheral functions, namely the maintenance, administration, and control functions, are highly complex, comprising many distinct tasks. Call control, for example, includes, among other functions, signaling interactions with users, database lookups for call treatment instructions, and service negotiations with the connection control function. The significance of the functional groupings presented in the model is not that they represent the fundamental elements of the network, but that, unlike the telephone network, these functions are indeed separable.

The functions of the network are both separable, in the sense that they can be provided with different equipment or from different locations, and modular, in the sense that not all functions are needed to support network services--they can be added or subtracted as a user sees fit. For example, one can simply contract the network provider for transmission channels, foregoing any transfer level functions. Virtual path layer services can be utilized without needing virtual channel switching capabilities. The different control modules can also be provided individually. A private network operator may own several VCI/VPI switches that contain sophisticated call control capabilities and thus would not subscribe to the call control functions offered by the network provider.

The model presented in Figure 6.2 refers only to the ATM platform network. As shown in Chapter 5, the platform network can be used to support a variety of communications applications. Some of these applications are better understood as networks. For example, the telephony and the connectionless data scenarios described above represent higher level networks. With a number of connectionless servers, one could construct a regional or even national connectionless data network that uses the ATM platform network transparently. To the user of such a connectionless service, the underlying ATM network would be effectively hidden. In relation to the model of Figure 6.2, such higher level networks may reside at many different layers. For example, in the third telephony application scenario described in Chapter 5 (Figure 5.3), telephone service uses network transmission and transfer functions and the service is provided entirely from the call control module, whereas in the second telephony scenario (Figure 5.2), higher level switching (using POTS central office switches) takes place on top of the virtual path transfer layer (virtual channel switching is not used). One could also construct networks that use platforms of only transmission services, foregoing any virtual path routing, or, in the case where private networks are connected with dark fiber,[57] networks that use only the physical medium as a platform.

The inherent separability and modularity of an ATM-based network affords tremendous flexibility to the providers of communication services. The separability of the functions enables an open, multi-vendor approach to network design. The modularity of the network allows for users and service providers to develop their own innovative network solutions, free from the traditional constraint of the availability of network features. The re-conceptualization of the network into these functional modules marks a break from past, when services were roughly categorized into switched services and pure transmission, or private line, services. Switching and transmission were taken as the basic units of the network, on top of which came enhanced call processing features. In this new environment, switching and transmission have been broken down into smaller, more fundamental functional modules. This deconstruction of the network enables the telecommunications community to halt the evolution of the network along the particular path that originated with the telephone system and start with a fresh approach to broadband communications.

---

[57] The term "dark fiber" is used in this context to refer to fiber that is provided to a customer without any transmission equipment to regulate the available bandwidth.

Chapter Seven

The Prospects for a Public ATM Network

7.1 INTRODUCTION

As the public telephone network becomes a public broadband communications network, it is not at all clear that it will naturally evolve into a ubiquitous, switched broadband network. In fact, as broadband capabilities have been introduced into the current network, they have not, by and large, been available on a public switched basis. High-speed leased lines are used to connect private networks. Frame Relay services, described in Chapter 2, are currently available on a semipermanent virtual basis only—there is no standard as yet for the signaling necessary to establish individual calls (Cavanagh, 1992). It is a premise of this thesis, however, that the fostering of a public, switched, broadband communications network is a long-term policy goal. The evolution of the public switched telephone network has proven the value of ubiquitous connectivity and, as people develop new, richer forms of communication by supplementing voice conversation with additional media, it is desirable that this connectivity—the ability to reach other subscribers through switched, public facilities—be maintained.

ATM switching offers the promise of forming the basis for a simple, universal, switched broadband network. Its flexibility in accommodating a wide range of applications and its scalability to support higher and higher data rates suggest that ATM technology may be sufficient to meet most broadband communications needs for quite some time. A public broadband network based on ATM could thus provide the stability necessary to foster the long-term development of new communications applications. This chapter examines the prospects for a public ATM network, comparing it with today's public switched telephone network, identifying a number of key implications of ATM networks, and examining ATM in the context of competing approaches to broadband networking.

7.2 THE ANALOGY TO THE PUBLIC SWITCHED TELEPHONE NETWORK

At first glance, a public ATM network seems vastly different from the public switched telephone network (PSTN). The interface speeds in ATM are orders of magnitude faster, the fundamental switching paradigm has changed, and the number of services supported has grown considerably. The changes are significant, but similarities between the two networks can also be found.

The first and most obvious difference between the two networks is the greater capability of the ATM network. A user of the PSTN, even when equipped with a highly sophisticated modem, is limited to communication at a rate of about 19.2 Kbps, whereas the CCITT has specified interfaces for ATM at 155 Mbps and 622 Mbps. The increase is staggering, a difference of four orders of magnitude. Services limited by the channel capacity of the PSTN can, with ATM, blossom into their full potential. Videotex services, up till now limited to ASCII text-based or crude graphical interfaces, could, over an ATM network, be presented with high resolution graphics and moving images. This increase in the maximum speeds supported by the network is accompanied by an increase in the flexibility of channel provisioning. Unlike the PSTN, where all services have to be adapted to fit into fixed rate voice channels, the ATM network supports a continuous range of channels, effectively doing away with the concept of the channel itself. This flexibility allows for a broad range of services and loosens the restrictions on users and service providers, enabling them to design services to run at speeds appropriate to the applications, not the speeds required by the network.

A second major difference between the old and new networks is the modularity of the network functions. As presented in Chapter 6, the network switching function has been broken down into separate building blocks, allowing for a separation of the transfer and control functions. In the PSTN, the network provider controlled all aspects of switching and all intelligent processing associated with the establishment and routing of telephone calls. The modularity of functions in an ATM-based network provides the opportunity to move much of the control of both the usage of the network and the design of new services into the hands of the users of the public network.

The third significant difference between the PSTN and a public ATM network is in the design approach. The PSTN was designed for the sole purpose of switching and transporting telephone calls. Channels were thus dimensioned to allow transmission of the human voice, no more and no less. The switching paradigm—circuit switching—was well-suited to voice communications and that it was non-optimal for data communications was irrelevant, for the network was designed to carry voice calls. The ATM network, by contrast, has not been chosen to support a single service, such as telephony, but to support the totality of communications services. A switching paradigm was chosen that would provide the flexibility to support the transfer of virtually any information that can be digitized. This move, from narrow optimization to broad generalization, is fundamental to the differences between the PSTN and the future ATM-based networks.

Despite, these basic differences, however, one can find equally fundamental similarities between the public switched telephone network and a public network based on ATM switching. First, like the PSTN and unlike networks such as cable television distribution networks, the public ATM network is designed as a switched, user-to-user network. In principle, just as is the case with the PSTN, any user in the country should be able to reach any other user in the country, even the world, for direct, end-to-end communication. While many of the discussions of broadband services center around video programming and information services, where interaction is between a user and a central source, one must not ignore the opportunities for user-to-user applications such as desktop videoconferencing or video telephony. By linking any user to any other user, a public ATM network extends the many-to-many principle of telephony to the broadband environment, and, as such, provides a unique communications service.

Although an ATM network is capable of supporting a far broader range of services than is the PSTN, the service independence of the ATM network is not new. The PSTN, after the *Carterfone* ruling, is similarly service-independent. Once AT&T relinquished control of what could be connected to the public network, it lost control over the services provided over the network. Voice traffic has been thus indistinguishable from data traffic and applications such as facsimile have flourished. This opening of the network interfaces, allowing users to control the type of information sent over the networks, set a basic precedent, limiting the role of the network provider to that of an indifferent carrier of electrical (or optical) signals, that will continue with the introduction of ATM-based networks.

A final, key principle held in common by the PSTN and a public ATM network is the role of the network as a fundamental platform upon which communication services can be provided. Facsimile communications and, to a lesser extent, packet switching services, were made possible by the presence of a standard, switched communications network—the PSTN. In early 1992, AT&T announced a new video telephone product that interfaces, not to digital, high-speed networks, but to the PSTN.[58] While the company could have chosen to focus its efforts on developing products for use with its own emerging network services such as N-ISDN or Frame Relay, it opted to stick with the low-tech network in place. AT&T's choice reflects the power of the PSTN as a fundamental, ubiquitous platform for switched communications. Note that this power of the PSTN comes despite its limited capabilities. The current evolution of the PSTN to support advanced network services such as N-ISDN and Frame Relay has led to something of a hodgepodge of network capabilities, serving different niche applications. ATM, because it is

---

[58] See "Consumer Videophone By AT&T," *The New York Times*, January 7, 1992, p. D1, col. 6.

designed as a total network architecture, not simply a new technique for relaying data, has the potential to play an analogous role to the PSTN. A public ATM network could serve as the common denominator, a network that may not be optimal for every application but that connects all users with standard, open interfaces. Innovative applications would be encouraged by the universality and the stability of the network. Although this notion of a single, fundamental, switched broadband network appears to be in opposition to the positions of those, such as Noam (1989), who argue for pluralism in networking, the ideas are not incompatible. A public ATM network would not serve to compete with the many, independently provided, application-specific networks so much as to form the platform upon which they are operated. In this vision, the ATM network supports the interconnection of the nodes of higher level networks and provides a standard interface through which users can reach these higher-level networks in the same manner as the PSTN enabled access to packet-switched networks.

7.3 IMPLICATIONS OF AN ATM-BASED PUBLIC NETWORK

The following discussions of public ATM networks, in this and the following chapter, are premised on a particular view of how these networks shall develop. Specifically, it is assumed that many of the network functions described in Chapter 6 will be available on a modular basis. Customers shall be able to lease dedicated transmission lines or semipermanent virtual paths. Virtual channel transfer services, presumably bundled with connection control functions, shall be available from the network provider. Finally, it is assumed that a basic set of call control services will also be available from the network provider, along with numerous specialized or enhanced call control services from independent service providers.

The public ATM network, in this view, is assumed to be an interconnected network of networks, managed and operated by multiple providers. To the network user, however, it should function, like the PSTN, as a single, seamless network. In the following discussions, the term "network provider" is used to denote the carrier providing the basic transmission and transfer services that form the platform for higher-level communications services.[59] This designation does not necessarily imply that these basic network services will be offered by a single, monopolistic provider—it is assumed, however, that even for highly competitive services, customers will have access to a carrier of last resort. Furthermore, linking the network provider to transmission and transfer services is not meant to preclude it from offering control or higher level services.

With this view of the future public ATM network, one can identify a number of qualities that will affect both its usage and its regulation. Lower transmission cost will reduce the effect of distance in network design and the flexibility offered by virtual path capability should spur the development of virtual private networks. The increased complexity of call control will enable novel forms of communication, thus frustrating efforts to regulate the provision of traditional services. Finally, the new paradigm of information transfer will mandate new approaches to pricing network services.

7.3.1 The Role of Distance

Improvements in optical transmission techniques continue to increase the transmission capacity of the existing network of optical fibers. The same pair of fibers that, when buried, carried 45 Megabits (10^6) per second, can now support SONET-based transmission at 2.4 Gigabits (10^9) per second. Laboratory experiments are now yielding speeds of Terabits (10^{12}) per second and few can pinpoint the ceiling of this progression. The effect of this ever-increasing capacity is a significant reduction in the cost of transmitting bits.

A classic problem for a network designer is the trade-off between the cost of switching equipment and the cost of the transmission links between the switches. Huber (1987) argued in

---

[59] This distinction, between the network provider and independent service providers, is not unlike the Japanese classifications of Type I and Type II carriers: Type I carriers provide services over their own facilities whereas Type II carriers offer services using facilities leased from Type I carriers (Sato and Stevenson, 1989).

The Geodesic Network, that as switching costs dropped relative to transmission costs, networks would be organized with the emphasis on small distributed switches. However, as Sirbu (1992) and Cheung (1992) have noted, this trend has been reversed—transmission costs are dropping relative to switching costs.[60] As transmission costs decline, distance is less of a factor, and switching, or other processing services, can take place remotely. Furthermore, ATM techniques such as statistical multiplexing and cross-connecting using virtual paths allow for more efficient usage of the transmission network, further encouraging the remote provision of services. A consequence of the possibility to provide services remotely is the reduction of barriers to entry for the provision of local exchange services. In this environment, higher level services, such as virtual channel transfer or connectionless data services, can be provided from fewer network nodes than the many central office switches used today for telephony. The changing economics of transmission are such that interexchange carriers, with their points-of-presence (POPs) in many local markets, could, if they had access to the access and transport services of local exchange carriers, offer many local exchange services.

As transmission costs diminish, distance begins to play a smaller and smaller role in the design of telecommunications networks. However, as Kleinrock (1992) notes, for some applications, distance becomes increasingly important. For a number of high-speed, real-time applications, the time for a bit to travel from one point to another in the network, known as the latency of the network, becomes a significant factor. Although the limits of the transmission capacity of an optical fiber are not well understood, the limit on the speed of transmission is well-defined; it is the speed of light. The system response to a user's command is therefore ultimately limited by the physical distance between the user's terminal and the remote resource with which he or she is interacting. For applications requiring very high data rates, this limitation can be problematic. Thus while the reduction in transmission costs will change the overall economics of

---

[60] Cheung shows a drop in transmission costs, between 1975 and 1990 (per Mbps • km) of greater than a factor of 100.

network design and service provision, leading to the remote provision of many communication services, such an architecture is clearly not suited to all services.

7.3.2 Virtual Private Networks

The capabilities and the modularity of an ATM-based platform network should lead to increased use of virtual private networks. In this discussion, a virtual private network is understood to be one in which shared public facilities are used to interconnect privately owned switching or processing nodes. Today, a private network usually involves the linking of customer nodes with either public, carrier-provided, dedicated lines or with private facilities, such as microwave links. The ATM platform network, by offering shared capacity through virtual paths, can obviate the need for dedicated links. Virtual private lines, using the virtual path concept, can be dimensioned so as to exploit the statistical nature of many communications and thus not to tie up the same amount of network capacity as dedicated private lines. This efficiency presumably leads to a reduction in cost, that, when combined with the overall reductions in transmission costs, should make the use of the public network an attractive option vis-a-vis the construction of private facilities.

The increased modularity of the public network also creates opportunities for virtual private networks. If switching functions can be unbundled along the functions introduced in Chapter 6, then operators of virtual private networks will have an increased ability to customize their networks by selecting different functions from different service providers. For example, a customer may contract for transmission, transfer, and connection control from a network carrier, but choose call control and higher level services, such as multimedia conferencing and connectionless data, from competitive service providers. The private network manager weaves these different functions together to create a customized, virtual private network.

7.3.3 Distinguishing Services in a Broadband Network

Distinctions between services such as telephony and television program distribution are central to the traditional regulation of communications in the United States. In a multimedia environment, with the bundling of different media into services under the control of countless service providers or even the end user, maintaining these and other distinctions between services will become increasingly difficult. The first challenge comes from the blurring of the lines that divide these traditional services. As Pepper (1988) notes, a videotext service, augmented to include moving images, begins to resemble a home shopping channel, which typically combines video with text and still images. Can one practically distinguish the two for regulatory purposes? A similar example would be an educational hypermedia service along the lines of a CD-ROM-based encyclopedia. Through this service, a child may read textual treatments of a certain subject, say a presidential election, and supplement this information with video clips of speeches of the candidates, activity on the convention floor, and television news reports of the campaign. Snippets of what was broadcast television are thus embedded within traditionally unregulated services.

A second obstacle to drawing distinctions between services is that the control of multimedia services is located at the edges of the network. The creation of a multimedia service, that is, the linking of different service components into a presentable service is a function of call control. As discussed earlier in the section of network functions, call control is not built-in to the network—it is better understood as a service that enables use of the network. As such, it is a highly flexible function and can be implemented in limitless ways. Services will not easily fall into distinct categories—they will, instead, be the products of the imaginations of the users and the service providers.

7.3.4 Broadband Pricing Issues

Telephone service has traditionally been priced with both access and usage components. A subscriber line, i.e. a connection to the public telephone network, is offered at a flat rate that usually includes a certain amount of local usage. Usage of long-distance facilities, and now, in some regions, local exchange facilities, is generally charged on a per-call basis. The rate structure for usage charges has been relatively simple in that the parameters associated with telephone calls are few, generally limited to duration, distance, and time-of-day. Beyond these considerations, a call is a call, occupying a single, well-defined fixed rate channel. "Calls" on an ATM-based network are significantly more complex—each is based on a negotiation for network resources. Parameters such as average and peak bit rates enter into every connection and each call is composed of multiple connections.

A number of approaches to usage-based pricing have been explored. Appleton (1991), speaking from the point of view of a network provider, recommends charges for both connection establishment and connection duration. The connection establishment charge serves to discourage users from establishing large numbers of very short connections and thus overloading the network's connection control function. Conversely, were there no charge associated with the duration of a connection, a user would have the incentive to establish a call and leave it up indefinitely, enabling it to function as a private line.

Establishing a usage charge based on the actual amount of traffic (i.e. the number of cells) generated has also been contemplated. This charge would serve to discourage users from wasteful uses of bandwidth and promote equity—heavy users would pay more than occasional users. However, the requirement of counting each individual cell associated with each connection could prove to be an expensive proposition and thus lead to cost inefficiencies. An ironic result would be the discovery that the cost of counting and billing for each cell switched was greater than the cost of actually switching the cells!

An alternative to per-cell charging would be to base pricing on the network resources allocated in the service negotiation performed at the time of connection establishment. In this approach, a customer would be charged for the capacity, that is the peak and average bit rates supported, of the virtual connection. This approach has the advantage of accounting for the opportunity cost of reserving network capacity. If a customer, through the service negotiation, reserves an average rate of 10 Mbps, but only uses 2 Mbps on average, the remaining 8 Mbps is wasted to the extent that the network, in allocating resources to other users, must plan for the full 10 Mbps. The disadvantage of this method is the difficulty of enforcement and the customer's incentive to under-declare expected usage if the customer believes that he or she can "get away with it." This point raises the issue of policing and punitive pricing. First, if the customer can generate more traffic than had been agreed and all of it manages to pass through the network, the customer is getting a certain amount of free service. If these excess cells are occasionally dropped in the network, and if the customer has an application, such as voice, that can tolerate a certain amount of cell loss, then the incentive to under-declare remains. If, however, all excess cells are dropped and fines are established for violations of service negotiations, the customer has a strong incentive to stay within the declared limits. Of course if the fines are too steep, then the incentive is for the customer to over-declare, and thus waste network resources.

The complexities of usage charging schemes and the rather low processing cost associated with cell switching has led others, notably Anania and Solomon (1988), to propose the elimination of usage charging, to be replaced by a single flat rate based on the bit rate supported at the access interface. This idea has received some support from Pepper (1988), who points out that rates based on utilized bandwidth will wreak havoc on prices for existing services. If a telephone call is priced at a penny a minute, suggests Pepper, a two-hour movie would cost over $800. This example, although it assumes that the cost of providing service is directly proportional to the bit rate utilized, highlights the constraints of marketplace realities—rational as it would seem to charge video programming providers and telephone companies the same for the bits they send across the network, very few users will watch movies at $800 a show.

The difficulty in developing workable pricing schemes for ATM-based network services is also rooted in the lack of knowledge of network costs. Until ATM switches have been manufactured and implemented on a wide scale, until call and connection control functions are more precisely defined and their implementations envisioned, these costs will remain mysterious. Pricing schemes, insofar as they are related to these costs, will be similarly hard to envision.

7.4 ALTERNATIVE APPROACHES TO BROADBAND NETWORKS

The public switched telephone network will not become an ATM-based public broadband communications network overnight. The introduction of ATM technology will be gradual and shaped by both public policy and market forces. ATM is a total network solution, not designed to fit any one application in particular, but it will face competition from networks that are optimized to meet the requirements of particular applications. In order to understand the prospects for a public ATM network, it is useful to examine some of the alternative approaches to broadband information transfer.

7.4.1 Frame Relay and SMDS

ATM is often mentioned in same breath as Frame Relay and SMDS as one of a series of fast-packet switching techniques that will contend for public data communications traffic. Frame Relay has the advantage that it is available today and thus will have a two to three year head start over ATM switching. While initially defined for a maximum access speed of 1.5 Mbps, Frame Relay is expected to support 50 Mbps eventually and laboratory experiments have yielded 100 Mbps performance (Zerbiec, 1992). This scalability suggests that Frame Relay, if it catches on, could serve the needs of many users for a number of years. Ultimately, as demand for access speeds reaches the B-ISDN levels of 150 and 600 Mbps, Frame Relay would likely give way to ATM. In the meantime, the two techniques may compete for certain applications of data transfer.

SMDS, scheduled to debut in late 1992, will follow Frame Relay into the market and also benefit from a head start over ATM. A key difference between SMDS and ATM is that SMDS is a connectionless service, and thus holds an advantage in the relaying of bursty data. Like Frame Relay, SMDS, should it gain acceptance, will pose a competitive threat to ATM in the market for high-speed data transmission. User investments in Frame Relay and SMDS equipment will be hard to dislodge, and ATM, arriving late on the scene, may have missed much of the market. To view ATM as simply another technological alternative for wide-area data networking, however, is to miss the basic point of ATM. ATM is an overall network solution, a platform upon which all forms of information transfer can take place. One should expect to see, even with widespread implementations of Frame Relay and SMDS services, ATM used as an all-purpose interconnection method. Nodes in both public and private Frame Relay networks may be connected using ATM (Figure 7.1), as could multiple SMDS switches (Figure 5.4, where the connectionless servers represent SMDS switches). ATM may also be used at the user interface, aggregating various of types of traffic for transport and transfer to different services across the network (Figure 7.2). Another possibility, suggested by Ali (1992), is the use of Frame Relay as an access ramp to ATM-based services (Figure 7.3). This view of ATM suggests a possible coexistence of ATM, Frame Relay, and SMDS, highlighting the complementary nature of ATM as an overall network platform service. One can also imagine, insofar as ATM is flexible enough to offer data communications services comparable to Frame Relay and SMDS, that users may find it redundant to transport, for example, Frame Relay traffic over an ATM network when the traffic could also be carried directly by the ATM network. One can thus expect, in the long run, a certain amount of cannibalization of Frame Relay and SMDS service by direct ATM transfer.

Figure 7.1.
Frame Relay switches interconnected by an ATM network.

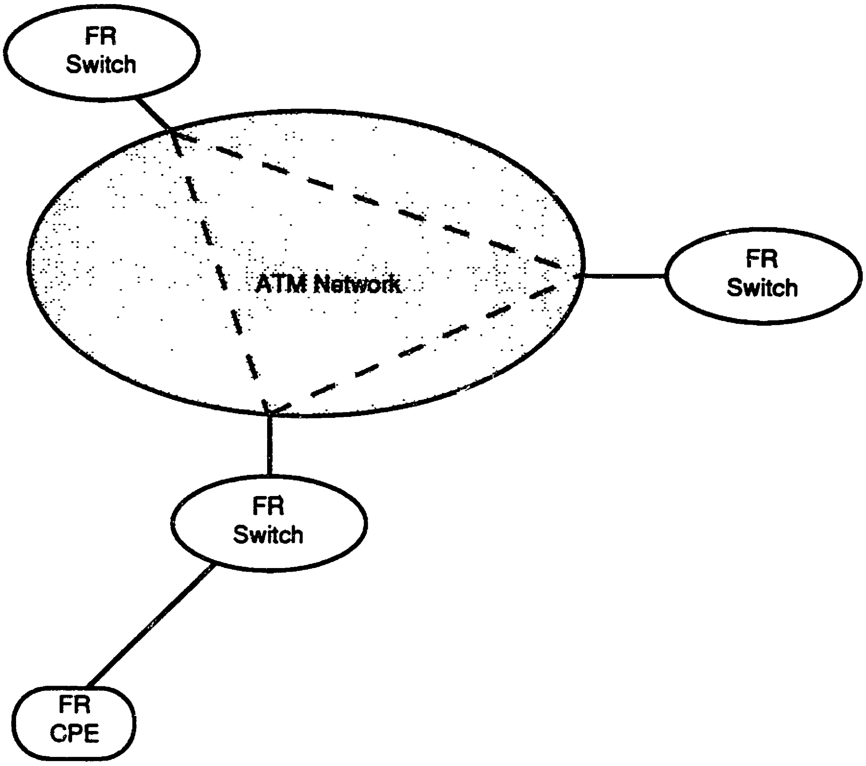


Figure 7.2.
ATM as an integrated on-ramp, providing access to different communication services.

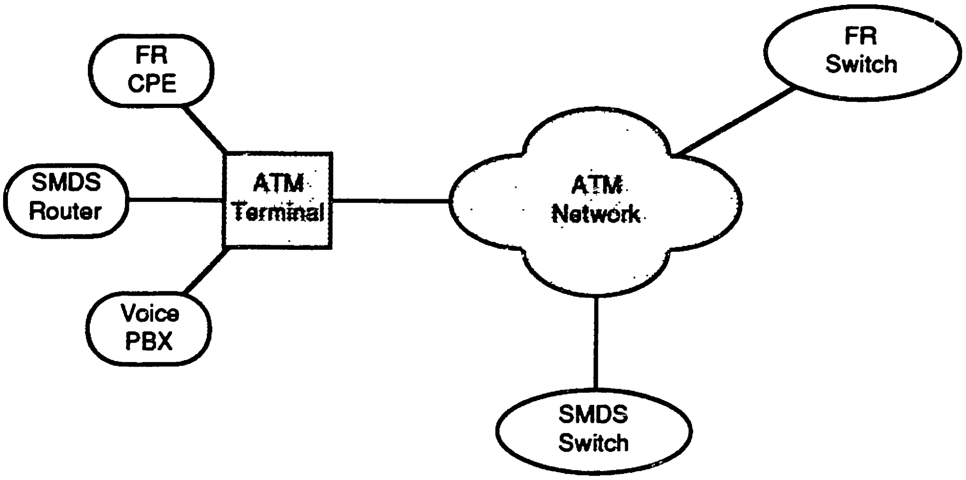
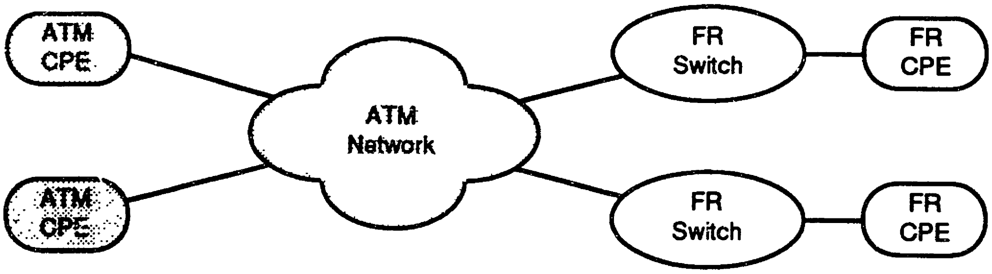


Figure 7.3.
Frame Relay providing access to an ATM network.



After Ali (1992).

7.4.2 The National Research and Education Network

The National Research and Education Network (NREN) was formally established in November, 1991 by the passage of S. 272, the High Performance Computing Act of 1991. The NREN is often touted as a national information superhighway, initially connecting university researchers to remote supercomputers and, ultimately, schoolchildren to the Library of Congress. In this view, the NREN seems very much like a broadband infrastructure, parallel to, and possibly competing with an ATM-based public broadband communications network.

From a technological and organizational standpoint, the NREN will resemble its ancestor, the Internet. The Internet, as defined in functional terms by Kahin (1992), is an international set of interconnected and interoperating autonomous networks that support a minimum of common functions. These functions, electronic mail exchange, file transfer, and remote log-in, are made possible by the standard use of the Internet Protocol (IP). The various local networks comprised by the Internet (there are over 5000 of them) are connected, in the United States, by NSFNET. As of 1992, NSFNET consists of a high-speed (45 Mbps) backbone network that connects dozens of regional mid-level networks. Access to these mid-level networks generally comes through leased circuits, either at 56 Kbps or at 1.5 Mbps (T1).

The NREN will, like the Internet, most likely utilize connectionless packet switches interconnected by leased facilities. In fact, the High Performance Computing Act mandates that the network be developed by purchasing "standard commercial transmission and network services" whenever feasible.[61] Whether these standard commercial services include ATM transfer services, or if dedicated transmission facilities will be preferred, is a matter for the designers of the network. In either case, it is important to view the NREN as a higher-level network, operating on top of a more fundamental transmission or ATM transfer network. The NREN is not so much a competitive alternative to a public ATM network so much as an

---

[61] 1991 High Performance Computing Act, § 102 (c)(8).

information infrastructure that will, along with other higher-level networks, enable productive use of a public ATM network.

7.4.3 Alternative Local Transport Providers

The local telephone monopoly has been challenged in recent years by the emergence of alternative local transport providers. These companies, most notably Teleport Communications and Metropolitan Fiber Systems (MFS), have deployed fiber optic networks in major U.S. cities, competing with the Bells and other local exchange carriers (LECs) for the carriage of high-speed data and telephone trunk circuits. The challengers generally offer point-to-point transmission services between customer locations, from a customer to an interexchange carrier's POP, or between POPs in a metropolitan area (Teske and Gebosky, 1991). Circuits are provided as a customer's primary connection or as a backup, or alternate route to the LEC-provided circuits. Although services were initially limited to dedicated lines, Teleport petitioned the Illinois Commerce Commission in 1991 to provide switched services in the Chicago metropolitan area (Burkhart, 1992). It may be expected that alternative local transport providers will attempt to enter additional markets in the future.

A number of regulatory decisions, by the state commissions in New York and Illinois and by the FCC, mandating that the LECs offer alternative providers equitable interconnection to their central offices, have promoted this competition in the local exchange. This regulatory support, combined with the advantages alternative providers are demonstrating in flexibility, responsiveness, and sometimes price, suggest that competition for some local exchange services is likely to intensify. As the new carriers become more established, more and more metropolitan customers will have multiple access and transport networks from which to choose.

If the incumbent local exchange carriers are so challenged, one might wonder if the acceptance of ATM in the public network would be jeopardized—will high-volume, sophisticated customers shun LEC-provided ATM services in favor of other transport and transfer offerings from alternative providers? The brief history of Teleport and MFS suggests otherwise. Both of these carriers have established themselves by providing industry-standard transmission services—T1 or DS-3 circuits—at lower rates or with better service. With the exception of providing dark fiber to some customers for point-to-point connections, their approach has been to compete directly with services also available from the LECs. Furthermore, the bulk of their business thus far has been in the interconnection of customers to local central offices or to inter-exchange carrier POPs. Most of their legal actions have concerned interconnection rights. Therefore, if their strategy is continue in this direction, alternative providers have strong incentives to meet the standard interfaces of the telephone network. If ATM is introduced into the public network with any success, and high-end users begin to connect their wide-area networks with ATM services, alternative providers will have to compete by offering their own ATM capabilities. One may also see, given that the alternative providers generally seem to move more quickly than the incumbent LECs, that, if high-end users show interest in ATM services in advance of LEC offerings, the alternative providers may be first to the market with local ATM services.

7.4.4 Cable Television Networks

The broadband networks of the telephone companies will face formidable competition from cable television networks. Cable operators already have a broadband infrastructure that can reach 80-90% of all U.S. households and are fast installing optical fiber in the trunks of their local distribution networks. While many telephone company executives may believe that the telcos are most capable of delivering home entertainment services, Pepper (1991) provides a number of persuasive arguments to show that cable television operators may indeed have a bright future. Cable networks can be upgraded to bring fiber optic quality television to the home for a fraction of the cost necessary to upgrade the telephone networks; cable operators are the incumbents, with nearly two-thirds of all households subscribing to cable service today; and cable operators have experience in packaging and marketing television programming, a completely new field for the LECs.

Cable television networks have been designed around the sole purpose of delivering one-way television programming to its customers. Up till now they have been provided with a certain amount of support for interactivity, allowing users to communicate with the network, but not with each other. Cable networks are one-to-many, or many-to-one networks, but not many-to-many networks. Most evolution scenarios (see Chiddix, 1990, for example) show the cable network continuing along this paradigm. However, the introduction of optical fiber offers new opportunities for cable television networks.

As optical fiber enters the cable network, initially for the purposes of increasing picture quality (by eliminating many of the amplifiers required for coaxial transmission) and offering more television channels, the capacity of the network can increase dramatically. While some of this capacity can be used to carry more channels, the rest can be used for other purposes, such as the carriage of telephone or other telecommunications traffic. Just as telco fiber residential video networks could multiplex B-ISDN services on the same fiber as video programming channels, cable networks can similarly segregate services over the same fiber infrastructure. Cable operators could compete directly with the LECs for local access and transport services in the same manner as Teleport and MFS. This competition has, to a certain extent, already begun. Cox Enterprises, one of the nation's largest cable MSOs, owns a 12.5% stake in Teleport.[62] Cable operators are also active in the burgeoning personal communications services (PCS) industry, having been granted various experimental licenses to test PCS systems.[63]

---

[62] On March 6, 1992, the Wall Street Journal reported that Cox was exercising its option to buy an additional 37.6% of Teleport. Cox's offer, which would raise its stake in ownership to 50.1%, came shortly after another large MSO, Tele-Communications, Inc. (TCI), purchased 49.9% of Teleport. Approval for both sales is pending.

[63] See Appendix C in U.S. Congress, Office of Technology Assessment (1991) for a list of experimental PCS licenses granted by the FCC.

Cable television networks will offer strong competition in the market for residential video services and in this respect, probably slow the adoption of residential B-ISDN. In the market for local telecommunications access, transport, and possibly transfer services, cable operators may also pose a challenge. In this latter arena, however, cable networks will appear very much like alternative local transport providers, and, if they are to compete directly with telephone companies and offer interconnection to long-distance networks, they will likely offer industry-standard interfaces. One of the properties of ATM technology is that it is not owned by the telephone companies—once the standards are in place and products available, any network provider can use them. Edward Horowitz (1992) of Viacom noted this insight at a recent panel discussion on digital television, stating that nothing was to prevent the cable operators from using the same equipment in their networks as is found in the telephone networks. The outlook for ATM, then, in the face of local exchange competition from the cable television networks, is based on its overall success in local, long-distance, public, and private networks. If ATM is accepted by users, service providers, and network providers, then cable television operators, if they are to compete in the provision local telecommunications services, can and will offer ATM network services.

7.4.5 Telephone Company Provision of Cable Television Service

Much of the discussion concerning telephone company provision of broadband network services tends to center around cable television services. Can the telephone companies provide better cable service at lower cost than current cable MSOs? Will the regulatory restraints that prevent telephone company ownership of in-franchise cable systems be lifted?[64] Will the telco-provided cable service be interactive? Cable television service is considered the "killer

---

[64] The FCC has decided on a video "dialtone" solution, in which telephone companies may carry video services over their networks on a common carrier basis, and has recommended that Congress lift the restriction prohibiting the direct provision of video programming by local exchange carriers. See In the Matter of Telephone Company-Cable Television Cross-Ownership Rules, Sections 63.54 - 63.58, *Further Notice of Proposed Rulemaking, First Report and Order, and Second Further Notice of Inquiry,* 7 FCC Rcd 300 (1992), and "F.C.C. Approves TV on Phone Lines," *New York Times,* July 17, 1992, p. D3, col. 4.

application" that will bring fiber to the home (or at least to the curb).[65] If the telephone companies are ultimately allowed to offer cable television services they will have a number of technical options for the delivery of video signals. ATM, of course, is capable of supporting the distribution services (i.e. broadcast television programming) envisioned for B-ISDN, but it is not necessarily the most efficient method of delivery. Television programs are, as they are encoded today, fixed rate sources, and, in general, do not require any switching except at the television itself. Thus the advantages of ATM are of questionable use to the problem of television program delivery and the processing and bandwidth overheads associated with the segmentation of video signals into cells may result in unacceptable inefficiencies.

Telephone companies are more likely to offer cable television services in a simpler, less expensive manner such as the passive, analog transmission of television signals. Although the development of telco-provided cable television services is almost entirely dependent on policy outcomes, and it is probably imprudent to speculate on those outcomes, it is quite possible that television programming will precede residential B-ISDN, or ATM-based, services. Thus fiber networks in the subscriber loop will carry cable television channels, possibly telephone service, and could eventually be upgraded to provide ATM-based services. This migration scenario is envisioned in Bellcore's 1990 Architecture Summary Report for fiber in the loop,[66] where B-ISDN and AM video are provided over the same fiber, but at separate wavelengths. This sort of coexistence, between ATM and residential video services, on the same fiber suggest the need for a unbundling, even if only for accounting purposes, of the fiber medium.[67]

ATM, in the long run, may serve as the vehicle for new types of video services, such as video-on-demand, or hypermedia programming, but, for the reasons described above, will

---

[65] See "Bell Atlantic Sees Television in Its Future," *Broadcasting*, May 8, 1989, p. 30.

[66] Bellcore (1990), discussed in Shumate and Snelling (1991) and Engineer (1991).

[67] The capacity of the fiber medium can be subdivided, not simply by bit rate, in the case of digital transmission, but also by wavelength. It is therefore not sufficient to develop models that assign costs on the basis of digital transmission capacity at a single wavelength. Instead, the cost for transmission at each wavelength must be identified; the costs for digital transmission can be derived from these more basic, per wavelength costs.

probably not, at least initially, find use in the residential market for the delivery of standard cable television programming. Passive video distribution systems and ATM-based services will, in all likelihood, share a common fiber network infrastructure.

7.5 CONCLUSIONS

ATM is clearly not the only solution for the delivery of broadband services and the telephone companies that created ATM will not be the only providers. To predict that the broadband world will be based on a single, ubiquitous ATM network is to ignore all of the forces that have led to multiple solutions in today's environment. It is equally unreasonable to view ATM as but one of a number of different data services or as a monolithic network itself. If one refers back to the conceptual model of the network, presented in Chapter 6, ATM is understood as a method by which information can be switched across a network. ATM forms a useful but not indispensable platform for the provision of higher-level services or can be used to connect users directly. Users, including service providers and managers of virtual private networks, of a public broadband network will have options. Some may connect their nodes with fixed rate transmission services while others may find use for different ATM functions, such as virtual path switching or full, call-by-call virtual channel switching. ATM is best viewed as a fundamental capability of the public network that will provide network users with powerful communications opportunities. That it is fundamental to the network, and standard across network boundaries, is pivotal to the success of a public ATM network.

ATM, because it is an inherently flexible and powerful technique, holds a certain appeal to designers of communications applications and communications equipment. This appeal is strongly enhanced by ATM's status as an international standard for public broadband networking. As AT&T Bell Laboratories executive Robert Lucky recently noted, certain user communities, such as the computer industry, may be dissatisfied with some of the characteristics of ATM, but have decided that if ATM is what will be available, then they will design products that use it (Lucky, 1992). Indeed, a number of local area networking hub manufacturers, such as Fibermux, SynOptics, Ungermann-Bass, and Cabletron, have announced ATM products, as have manufacturers of multiplexing equipment. The ATM Forum, an industry group composed mainly of manufacturers of telecommunications and local area networking products, has already settled upon a preliminary specification for an ATM user-to-network interface.[68]

ATM is now even discussed as a long-term solution for Gigabit local area networks (Kung, 1992). LANs have traditionally developed using the shared medium approach, in which users are connected to a ring or a bus, along which data packets for all users are transported. Each terminal receives all packets, but only "reads" packets containing its address. With the coming need for Gigabit capacity to the desktop, the shared medium approach is likely to give way to a switched-star architecture. The development of premise-based ATM networks would represent a new convergence of public and private architectures and facilitate the interconnection of geographically disparate networks. In the current network environment, the circuit-switched, fixed channel approach of the PSTN clashes with the shared medium, packet-based premises networks, leading to an inefficient mismatch of communication techniques. Acceptance of ATM techniques in premises networks naturally bodes well for future demand of public ATM services.

Given the push coming from the user and CPE manufacturing communities, one may expect demand for ATM connectivity to outpace the network carriers' development of ATM standards. The CCITT's progress on ATM standardization has been slow to date, and key components are not expected for a number of years. Specifications for signaling, necessary for the provision of call-by-call switching, are not expected until the late 1990s (Bonatti *et al*, 1991). These delays will likely result in a phased introduction of ATM capabilities. As users develop ATM interfaces to their local area networks, they may wish to have them connected with semipermanent virtual paths. Such cross-connecting services, scheduled to be available this year

---

[68] "ATM Forum Spec Done," *Communications Week*, June 22, 1992, p. 3, col. 5. For a discussion of the ATM Forum's activities, see Giancarlo (1992).

in France on an experimental basis (De Prycker *et al*, 1990), could surface in the U.S. in the next two years. In the years following, more sophisticated ATM services could be added in phases as made possible by progress in standardization and as encouraged by the development of new applications.

Acceptance of ATM technology in the user community does not guarantee widespread availability of public ATM services. Introduction of ATM switching in the public network will likely focus on private wide area networking. Whether early implementations of ATM will eventually lead to public switched ATM services may be decided by public policy as much as market forces. As noted above, however, this thesis takes as its premise that the United States will eventually have a public switched broadband network and that it will most likely be based on ATM technology. It is beyond the scope of this thesis to argue for government intervention to assure the availability of such a network. Rather, the purpose of the thesis is to identify key policy issues and opportunities that will arise should the network so develop. These issues are addressed in the following chapter.

Chapter Eight

Policy Recommendations

8.1 INTRODUCTION

In Chapter 5, it was shown that, with the widespread diffusion of ATM technology, the public telephone network would be fully transformed into a fundamental, service-independent communications network. Telephone service would become one of a number of possible services provided on top of this more basic platform network. It was further argued, in Chapter 6, that, with the reconception of network design that has followed the CCITT's selection of ATM for the architecture of future networks, the functions of a modern communications network are separable and modular. These two developments highlight the fundamental differences between future broadband networks and the current telephone network.

The assertion that the telephone network and its barely recognizable descendant, the public broadband communications network, are two fundamentally different beasts carries with it the implication that the two networks ought to receive different regulatory treatment. The focus of this thesis has been on the transformation of the public network brought about by the development of ATM technology. Constructing policy around a specific technological development is a hazardous venture, as policy and technology tend to move at dramatically different paces and obsolescence is a legitimate concern. The challenge for policy makers is to see beyond the details of a particular technology and recognize the more fundamental trends that accompany the technological development. Whereas it would be foolish to develop detailed policies around the specific properties and features of Asynchronous Transfer Mode, it would not be unreasonable to develop policy that reflects the qualitative changes in the network that will result from the adoption of ATM principles.

In examining this issue, one is also drawn to a chicken and egg question. The fundamental changes in the network are by no means necessarily the consequences of ATM technology—one could argue that the development of ATM is a response to the changing communications environment. The continual advances in optical fiber technology led to the capability to transmit signals of very high bandwidth, and out of the need for a switching platform fast enough and flexible enough to handle a broad range of signals ATM was born. Rather than creating change, then, ATM can also be viewed as the embodiment of a change and thus an examination of ATM serves to probe the dimensions and the nature of this change.

The development of a new switching platform for the public network also provides for policy opportunities. As the technological paradigm of the network is being reconsidered, the opportunity exists for policy makers and technologists to work in concert, or at the very least, to work with an awareness of each other's goals. In the case of developing both technology and policy for broadband networks one must draw a distinction between government policy attempting to direct the development of technology and government policy influencing the development of standards and architecture. In essence, the distinction is between technology and implementation. The government would not, for example, want to mandate the use of a particular multiplexing technique for optical fiber transmission. However, the government may, in another example, wish to provide guidelines for the unbundling of services and the opening of interfaces that would reflect general policy goals. These guidelines would have the effect of challenging network providers to develop standards and make architectural choices that facilitate the attainment of the government's policy goals.

Given that ATM-based communications networks will not become a reality for several years, given the uncertainty of their development, and recognizing the hazards of developing policy too closely around technological considerations, any policy initiatives in this area must be undertaken with caution. Nevertheless, one can envision a number of general policy directions that might be explored at this time.

8.2 THE TELEPHONE NETWORK AND THE UNDERLYING BROADBAND NETWORK ARE LOGICALLY SEPARATE

The analysis in Chapter 5 showed that the provision of telephone service is logically separate from the provision of the underlying network services that support telephony and other applications. Tomorrow's public network is not simply a telephone network with other, "enhanced" capabilities, it is a service-independent, all-purpose communications network, capable of providing a neutral platform upon which many different services can be offered. Telephone service, rather than being integrated into this basic platform, is one of the many higher-level services the platform supports. For the purposes of creating policy for future communications networks, it is useful, even imperative, to recognize this distinction.

The first practical reason for conceptually disaggregating telephone service from the provision of network infrastructure is that their respective economic characteristics are different. The history of regulation of the public telephone network has been justified by the consideration of telephone service as a natural monopoly. As policy makers began to recognize that certain services provided by the telephone network no longer appeared to justify their designations as natural monopolies, the network was subdivided, to the extent practicable, into monopoly and competitive components. Hence the AT&T divestiture accord, which severed long-distance telephony from local exchange services, and the *Computer II* proceedings, which drew the line between basic and enhanced services. Today, the viability of the natural monopoly in local exchange services is under much scrutiny, given the emergence of alternative local carriers and low-cost wireless technologies.

Part of the difficulty in discussing competition in local exchange services lies in the imprecision of the term "local exchange services." Challengers such as Teleport and MFS are not providing local exchange *telephone* services, per se, for their focus has been largely on the aggregation of telephone traffic over high-speed private lines. Wireless PCN services, while certainly not limited to telephony, are currently envisioned as narrowband, low data rate services. Wireless PCN technology has been heralded as the vehicle for breaking the local exchange monopoly, but if one is to characterize it as such, one must be more specific. If PCN-based telephone service can be offered at a price comparable to conventional wireline telephone rates, then the natural monopoly of local exchange *telephone* service may no longer be viable and we may enter a new era of competition.

The end of the natural monopoly in local *telephone* service does not imply, however, the end of any natural monopoly in local exchange services generally. Provision of telephone service, be it wireless or wired, is a different engineering problem, requiring different technical solutions, than the provision of a high-speed, broadband communications network. While wireless implementations are possible, broadband networks are typically envisioned as optical fiber networks. The provision of optical fiber access involves costly digging, trenching, and the laying of cable, a far different matter than the construction of radio towers. The economics of providing telephone service are thus different from those involved in the provision of broadband networks.

In addition to having different economic characteristics, telephone service and broadband network provision may also be subject to different policy goals. For example, regulation of telephone service has long been governed by the principle of universal service. State regulators across the nation have made the determination that telephone service is so basic to human existence that all people should have access to it. In return for the legitimization of their monopoly status, telephone companies are required to offer service to even the most remote parts of their franchise areas. For years, long distance rates subsidized local rates so that access to the telephone network would be affordable to all. The universal service goals, while possibly non-optimal from an economic standpoint (Noam, 1992), reflect a judgment that there is a public interest in the availability of telephone service that transcends market forces.

The telephone network is also considered vital for emergency communications and military preparedness; its reliability is therefore a matter of public policy.[69] These same reliability concerns may not necessarily apply to the underlying broadband network. First, one must recognize that telephone service is one of many different communication services to offered on top of a basic network platform. Not all of these communication services will require the same reliability characteristics as telephone service.[70] Moreover, the reliability concerns of these higher level services can be addressed by the service providers. Just as error correction in an ATM network is done largely at the edges of the network, by the user's or service provider's equipment, reliability can be designed into the higher level networks. Service providers can utilize diverse routing and employ redundant equipment to ensure the smooth operation of their networks in the event that some of the basic network components fail. This analysis is not intended to imply that the basic network need not be designed for a high degree of reliability. However, by providing a minimally acceptable level of reliability in the basic network and leaving the task of ensuring ultra-reliability to the users and service providers that require it, one does not force users of the network to pay for extra degrees of reliability that they do not value. The approach to network reliability may thus be different for telephone service and the underlying broadband network.

It is argued in this section that policy makers should make a conceptual distinction between the provision of telephone service and the provision of basic network services. The policy goals, such as universal service, that are associated with telephone service, are not necessarily the same as those that would be applied to the underlying, service-independent communications network. This observation does not imply that policy goals for the two networks are mutually exclusive nor that the goals set out for telephone service should not be

---

[69] During the AT&T antitrust deliberations, then Secretary of Defense Caspar Weinberger argued that the breakup of the Bell System would be detrimental to the national defense (Temin, 1987).

[70] A video dating service, for example, might not have the same reliability concerns as a military communications network., although other privacy and security issues might nevertheless be of concern.

extended to basic network service. Instead, attention is called to the special role of telephone service. What holds for telephone service does not automatically hold for basic network services—their roles in our society will be different. Policy goals should be established for service-independent communications networks with attention to their own special characteristics and these goals may or may not diverge from the goals that guide the regulation of telephone service.

8.2.1 A New Approach to Cost Allocation

The preceding paragraphs urge the conceptual, or intellectual separation of telephone service from the network that supports it. In addition to this abstract approach, concrete steps may also be taken. One of the continuing headaches of telecommunications regulation is the difficulty of allocating costs between regulated and non-regulated activities. This issue is further complicated by the introduction of new services into the network, and is especially problematic when network capacity can be dynamically reconfigured to provide completely different services (Pepper, 1988). The development of service-independent, broadband networks will not make this problem go away, but it may offer an opportunity for a new approach to the allocation of costs.

An ATM platform network is indifferent to and unaware of the services provided upon it. Discriminating among services is therefore problematic. A network provider cannot charge more for a given amount of capacity to be used for telephone service than if it were to be used for high-speed data communications because the network provider has virtually no way of knowing how the capacity is being used.[71] In order for the network provider to charge for use of the network, service-independent tariffs must be developed. This concept is not new, for today, a telephone call is priced the same whether it carries human voices or modem traffic, just as the price of a private line is not dependent on whether it carries voice, data, or video. However, the

---

[71] A network provider could, by analyzing the traffic characteristics of certain virtual channels or paths, possibly surmise the service associated with the traffic. For example, a steady stream of 64 Kbps implies a voice channel. As services begin to integrate different media, however, they will become harder and harder to define (see section 7.3.3), let alone identify.

cost parameters, such as distance and duration, for conventional telephone calls are fewer than for ATM-based calls, which would have additional cost components associated with data rates, burstiness, and service quality. Furthermore, if network services are to be provided on a unbundled basis, costs cannot be spread across vertically integrated functions—in order to limit arbitrage, the rates charged for individual functions must track their respective costs.

One of the problems of joint cost allocations is the linking of the telephone network to the enhanced services it can support. A frequently employed approach to this problem is to determine the incremental cost of adding an unregulated service to the telephone network and allocate this incremental cost to the enhanced service. While this approach does in theory shield the consumer of basic telephone service from additional costs due to the enhancement of the network, it does not necessarily reflect the value of the underlying network infrastructure to the provision of the new service. If, for example, an enhanced service can be provided by the addition of a small module to a large central office switch, the incremental cost of providing the enhanced service may be quite small. However, the telephone central office switch may well have been designed in a manner so as to facilitate the provision of the enhanced service with a simple add-on module. If the switch were not so designed, the incremental cost of providing the enhanced service might be considerably larger.

Rather than attempting to separate the costs associated with the enhanced service from the costs of providing telephone service, one could instead identify the service-independent components associated with basic telephone service and various enhanced services, and assign to them these costs in a non-discriminatory manner. For example, if a telephone service provider and a video service provider each need high-speed virtual paths to connect nodes in their networks, the challenge is to determine the costs to the underlying network provider of providing these virtual paths. These service-independent costs are then assessed, or charged, to the respective service providers. In this scheme, the investment in infrastructure, i.e. the burying of the fiber and the purchase of the ATM switches, no longer enters into the calculation of telephone rates directly. Telephone rates would be based on the prices, be they internal transfers or open market purchases, charged to the telephone service provider for the basic network services needed to support a telephone network. This approach could be viewed as an extension of Open Network Architecture. Under ONA, service providers build their "enhanced" services from a tariffed menu of basic service elements. Telephone service providers, under this new approach, would build their "basic"[72] service from a menu of service-independent components such as virtual paths, virtual channels, or dedicated transmission channels.

The cost allocation scheme presented in the preceding paragraph does not imply any sort of divestiture or structural separation between the telephone service provider and the underlying network provider. In the years ahead, during the transition to a fully service-independent broadband network, they will frequently be one and the same. As costs will be assessed internally, the scheme does not necessarily alleviate the problem of potential cross-subsidization among monopoly and competitive ventures. The competitive threats of both wireless telephony and alternative local transport may, however, provide some checks against price distortions. If, for example, a network provider were to overcharge for low capacity services in order to lower its costs for high speed services such as data communications and video delivery, it would risk having its telephone business undercut by wireless providers.

While widespread broadband, service-independent networks will take some time to emerge, this costing approach could be employed today on a smaller scale. Certain elements of the what is considered the public switched telephone network are service-independent and used for multiple purposes. Inter-office transmission facilities, for example, carry traffic for both public and private networks. One could begin to move toward service-independent costing today by developing a method for estimating the costs of providing transmission services (e.g. T1, DS-3, or SONET channels) over these existing facilities. Costs currently associated with these

---

[72] The term "basic" is used here to refer to the FCC's *Computer II* definition of "basic service" (section 3.3.3) as opposed to the basic transmission and transfer services discussed throughout this thesis.

facilities in the determination of a telephone company's overall costs could be removed and replaced by the newly determined, service-independent costs associated with the actual transmission capacity dedicated to telephone trunks. These latter costs would then be linked to the rates charged for identical services offered to private customers or service providers.

Adopting a cost allocation principle that disaggregates telephone service costs from underlying network costs may reveal subsidies to local users that are embedded in the current regulatory arrangements. Such a result does not require that consumers pay higher rates, however, as these rates can ultimately be adjusted through the political process. Any implicit subsidies would instead become explicit. As discussed earlier in this chapter, we may wish to adopt special policies for telephone service. If we are currently distorting the prices of telephone service for public policy reasons, we may wish to continue to do so. Under this new approach to cost allocations, we may have to subsidize local telephone rates explicitly if we are to maintain any existing subsidies. However, until we determine underlying network costs and assess them accordingly, we will not understand fully the extent of these subsidies, assuming that they do exist.

Implementation

A shift in cost determination and allocation principles is a daunting proposition and would need gradual implementation. The fact that local exchange carriers are, for the most part, regulated by the states, enables a certain amount of experimentation with respect to this initiative. Individual state utility commissions could examine this approach and gain experience with the practical difficulties of implementation. As more and more of the components of the telephone network become service-independent, their costs could be folded into this regulatory scheme. In this scenario, the fundamental change, of disaggregating telephone regulation from the regulation of basic network services, need not take place all at once, as through the passage of a law, but can evolve, as appropriate, over time.

Such a bifurcated approach to regulation does raise the issue of what Pool (1983) called the "convergence of modes." As the telephone network becomes capable of supporting services, such as broadcast video, that are traditionally provided over networks subject to vastly different regulatory schemes, which rules shall apply? If a network provider is supplying transmission and transfer services to both a telephone service provider and a video programming distributor, should the network provider be regulated as a telephone company, as a cable television operator, or as something else entirely? This question also raises a jurisdictional issue—telephone companies are regulated by the states whereas cable operators are franchised by local municipalities. This jurisdictional question is currently being discussed in the debate over the FCC's video "dialtone" proposal.[73] The video dialtone proposal contemplates the provision of video services over telephone company networks on a common carrier basis. The Commission has tentatively concluded that a local exchange carrier (LEC), under this arrangement, would not need local franchising authority under the 1984 Cable Act. Ultimately, as the boundaries between traditionally separate networks and services begin to blur, new regulatory policies will be needed. The move toward a bifurcated regulatory approach will not only underscore this need but may also provide insight into the challenges of multimedia services regulation.

Implementing a new, service-independent cost allocation scheme that draws a distinction between telephone service and basic network service will be challenging and arguably premature. Broadband, service-independent networks are years from becoming ubiquitous, and the public network, for all of its advances, still exists primarily for the purpose of providing telephone service. The principal reason for pushing ahead with a new regulatory approach is, quite simply, that the need to determine service-independent basic network costs is inevitable. Broadband, multi-service networks will be integrated into what was the public switched telephone network in due time and problems of cost allocations will remain.

---

[73] See In the Matter of Telephone Company-Cable Television Cross-Ownership Rules, Sections 63.54 - 63.58, *Further Notice of Proposed Rulemaking*, 7 FCC Rcd 300 (1992).

The introduction of competition into local exchange services, either telephony or broadband, will also exacerbate the need for accurate cost information. The rise of MCI and subsequent demise of the Bell System can be traced to a pricing system that did not reflect actual costs—an "unsustainable" rate structure (Phillips, 1991). As alternative local carriers challenge the LECs for network access and transport services, the LECs, if they are to compete fairly, will require the ability to charge according to the real costs of providing service. If PCN-based wireless service is to compete head-on with conventional telephone service, the rates associated with LEC-provided telephone service will similarly need to reflect actual costs.

Separating telephone service from basic network service has the additional advantage of allowing telephone service to evolve competitively. If wireless telephony can effectively compete with wired systems, the regulation of telephone service can be relaxed accordingly. Telephone service providers can then compete on the basis of sophisticated calling features that would confound distinctions between basic and enhanced services. The distinctions between different services will in any case become increasingly problematic as media are mixed on a call-by-call basis. Extracting telephone service, for regulatory purposes, out of an array of new multimedia services will become an intractable problem. Separating telephone service from basic network infrastructure services will enable the regulation of telephony to evolve independent of the regulation of the basic network. Given the differences in the economics and, possibly, in the policy goals, of the two networks, such a decoupling is most appropriate.

8.3 THE SEPARABILITY AND MODULARITY OF FUNCTIONS CREATES OPPORTUNITIES FOR NETWORK OPENNESS

The separability and modularity of the network functions described in Chapter 6 provides an opportunity to explore the possibilities for unbundling the network along fundamental, conceptual lines. Whereas much of the network unbundling witnessed today in the FCC's Open Network Architecture process consists of selling standard call features individually, rather than grouped in packages,[74] the current B-ISDN standards process is breaking the functions of the network into more fundamental building blocks. As the network becomes broadband and grows to accommodate a greater range of services, the FCC is faced with the decision of whether or not to extend the ONA concept and apply it the coming networks.

Open Network Architecture, as discussed in Chapter 3, was born in the FCC's Third Computer Inquiry. In that proceeding, the Commission sought to lift the structural barrier that had been erected in the Second Computer Inquiry between the BOCs' deregulated enhanced service operations and their provision of regulated, "basic" services.[75] The structural separation was replaced by non-structural safeguards, notably the Comparably Efficient Interconnection (CEI) and ONA requirements. CEI required that each BOC, when offering an enhanced service, file an interconnection plan that enabled all enhanced service providers offering the same service to have comparable access to the same features and functions to be used by the BOC. The Commission sought a way to avoid the burden of approving CEI plans for each new enhanced service offering and thus required the carriers to develop Open Network Architecture plans. Networks designed using ONA principles would, the Commission believed, foster equal access goals because the BOCs and other enhanced service providers would be constructing their enhanced services from the same menu of features and functions, available on comparable terms. A key element of ONA was the development of this menu—the "unbundling" of the network into "basic service 'building blocks'" known as "Basic Service Elements," or BSEs.[76]

---

[74] See In the Matter of Filing and Review of Open Network Architecture Plans, *Memorandum Opinion and Order,* 6 FCC Rcd 7646 (1991) for the most recent listing of Basic Service Elements provided by the BOCs.

[75] Third Computer Inquiry, *Report and Order,* 104 FCC 2d 964 (1986).

[76] Id. at 1064.

8.3.1 Current Difficulties with the ONA Approach

The brief history of ONA has been highly controversial, with the degree to which the BOCs have unbundled their networks being the major point of contention.[77] Enhanced service providers have, by and large, argued that the BOCs have not unbundled the network into fundamental building blocks, but have instead focused on providing supplemental, optional, features and functions.[78] The late William McGowan (1991), former chairman of MCI, once dismissed the BOCs ONA efforts out of hand, stating that

> "ONA as an innovative procompetitive tool is easily dismissed by anyone familiar with the relevant proceedings at the FCC. The RBOCs are simply putting old wine in new bottles. They have done little to meet the demands of their customers."

In responding to these complaints, the Commission has argued that to require a fundamental unbundling of the network, although a potentially valuable long-term goal, would be too costly, given the current configuration of the network. Therefore, it has required the BOCs to unbundle BSEs "only to the extent technologically feasible given the current network."[79] Whether or not one chooses to accept the BOCs' insistence that, at the time, fundamental unbundling was technologically infeasible, one cannot dispute the facts of the architecture of the network. The digital central office switches that are now widespread through the BOC networks were designed in the early 1980s, well before the ONA requirement was introduced. The location of the call processing intelligence is only now, with implementations of Intelligent Network concepts, being removed from the central office switches. One could argue that Open Network Architecture was, when introduced, something of misnomer. ONA did not induce a new architecture in the public network so much as require that services provided under the existing architecture be offered in an unbundled manner. The reluctance of the BOCs to comply with the ONA demands of the

---

[77] See In the Matter of Filing and Review of Open Network Architecture Plans, *Memorandum Opinion and Order*, 4 FCC Rcd 1 (1989), at 37-41, paras. 59-68 for a variety of comments of the issue of unbundling.

[78] Id. at 39, paras. 62-63.

[79] In the Matter of Intelligent Networks, *Notice of Inquiry*, 6 FCC Rcd 7256, (1991) at 7256, para. 2, hereafter *Intelligent Networks/NOI*.

enhanced service providers, while possibly a combination of strategic, economic, and technological considerations, has been a reluctance to design and implement a new network architecture to reflect the policy goals of ONA.

Because many of the disputes over unbundling have been complicated by the network architecture and technology that existed at the time of *Computer III*, the development of new network architectures and technologies offer, at least in principle, opportunities to address some of the issues of fundamental unbundling. The BOCs' development of the Intelligent Network concept and their current efforts to realize it in the form of Bellcore's Advanced Intelligent Network (AIN) recommendations,[80] appear to present such an opportunity. The Coalition of Open Network Architecture Parties (CONAP) petitioned the FCC to conduct an inquiry into the BOCs plans to upgrade their networks around an AIN architecture for the purpose of ensuring that the future network will be "modular and transparent."[81] The FCC has accordingly opened a docket, *In the Matter of Intelligent Networks*, in which it will examine "the interrelationship of Open Network Architecture (ONA) with emerging network design."[82] CONAP has expressed concern that the AIN envisioned by the BOCs will perpetuate the closed nature of today's networks.

The Intelligent Networks docket may serve as a test case in exploring the nature of the FCC's role in shaping the architecture of the public network. The FCC is in a position where it recognizes the dissatisfaction among enhanced service providers over the implementation of ONA and is now presented with an opportunity to influence the architecture of the network to reflect its policy goals. The Commission has tread cautiously on the issue of unbundling as a matter of principle, acknowledging that further unbundling "*may* be in the public interest."[83] The proceeding will give the Commission an opportunity to examine the opportunities for unbundling and to weigh the costs and benefits of different degrees of unbundling. The matter

---

[80] For an overview of AIN, see Berman and Brewster (1992).

[81] Intelligent Networks/NOI at 7257, para. 8.

[82] Id. at 7256, para. 1.

[83] Id. at 7256, para. 2, (emphasis added).

will also force the FCC to make a determination of the extent of its role in the development of architectural standards for the network. If, for example, the Commission should conclude that current AIN implementation plans do not adequately open the network for innovative competition in enhanced services, how would it impose openness requirements on the carriers?

8.3.2 Opportunities Created by the B-ISDN Standards Process

The requirement to develop signaling standards for B-ISDN that are capable of supporting a broad range of multimedia services has led to the two-track standards development process noted in Chapter 4. On one level, the signaling systems developed for Narrowband ISDN are being modified to provide enhanced capabilities. On another, fundamental communications concepts such as the call model (i.e. what constitutes a "call") are being revisited. It is important to recognize that as these concepts are being rethought and re-engineered, they are not strictly dependent on ATM technology. ATM is a method of implementing connections in the network whereas many of the call set-up and call control functions can be designed independent of the connection method employed by the network. Thus, as the telecommunications industry develops new approaches to defining and establishing calls within a network, these approaches are not wedded to the ultimate success of ATM as a transfer mode—they can presumably be applied to other modes as well.

The B-ISDN standards process represents a great opportunity to shape the development of future communications networks in the public interest. As noted earlier, the B-ISDN approach reflects a new paradigm of network design and, as it develops, holds the potential for even greater change. As communications principles are being examined at a very basic level, participants in the process have the opportunity to shape the direction of future communications capabilities. In particular, if the unbundling of the network into more fundamental, more basic building blocks than has been accomplished to date with ONA is deemed to be in the public interest, then the B-ISDN standards process offers an opportunity for such unbundling to be designed into the architecture of future networks.

In Chapter 6, a conceptual model of the functions of the network was presented. Implicit in the organization of the model is the separability of the identified functions. Transmission services can be provided with or without higher level ATM transfer services. Similarly, ATM virtual path routing could be offered without virtual channel switching. Functions that interact on multiple levels are similarly separable. The physical implementation of these functional modules is not pre-determined. They are organized as independent functions, and can, in principle, be provided by separate entities, so long as they communicate at standard interfaces. In particular, one can imagine the separation of call control from connection control. If standard signaling primitives can be agreed upon for communication between the two modules, there is no a priori reason why they must be integrated. The same logic holds true for the relationship between connection control and the administration module. The separability of these functions, as currently contemplated by the CCITT, lends itself to the principle of open architecture espoused by the FCC.

Much of the discussion of these functional modules has been in abstract terms and it is useful to examine how they would apply in the current telephone network. Today, a caller specifies a call by dialing the number of the target party. The call is processed based on the service logic associated with the telephone number. If, for example, the telephone number dialed is associated with a call forwarding service, the call is routed to the "forward to" line. The logic that governed the treatment of the call is embedded within the software of the terminating central office switch and thus both the call control and connection control functions are combined in the service. In a network where these functions are implemented separately, one may find a call control module that, after receiving the caller's request, determines appropriate routing treatment by checking various databases throughout a call control network, then requests a connection between two physical terminals (i.e. telephones) from the connection control module. In this implementation, the calling features of the network are kept entirely within the call control module. From an ONA perspective the difference is clear. In the former implementation, unbundling takes the form of separate calling features such as call forwarding. In the latter example, the entire call control function could be unbundled, leaving the actual design of features, as opposed to the combining of existing features, to the enhanced service provider.

The design of the network around basic conceptual functions provides an opportunity for truly open architecture. The argument that fundamental unbundling is made problematic by network technology will no longer be viable should the network develop in an open manner. From this point on, the degree to which broadband, intelligent networks will be open is a matter of policy. The systems concepts behind the new network lend themselves to openness, but physical implementations of these concepts may effectively close the network. The network providers and the manufacturers of network equipment can choose integrated solutions, where conceptual functions are combined within hardware and software, thus limiting access to some of the basic functions of the network. The policy community is presented with an opportunity to examine the possibilities for fundamental unbundling of broadband networks and to determine the ultimate openness of the network.

As Reed (1991) has pointed out, unbundling is not without its costs. These costs take two forms. First, if there are actual economies of scale and scope between elements of the networks that are unbundled, inefficiency costs are incurred. Second, opening the network raises security concerns and may require more robust, and hence more costly, interfaces. The most compelling argument for network openness is that an open platform can support innovative services from a multitude of entrepreneurial service providers and lead to a greater utilization of the network's capabilities. This argument was summarized by Mitchell Kapor of the Electronic Frontier Foundation at the FCC's 1991 hearing on "Networks of the Future":

> " Just as the Apple II personal computer was a platform that allowed others to invent new applications, the [National Public Network] can be a platform for information entrepreneurship."[84]

Openness, to the extent that it promotes competition and innovation, is clearly desirable. Determining the proper degree of unbundling, however, is far from trivial, for any effort to open the network must strike a balance between the costs and benefits of such openness.

The FCC's inquiry into the Advanced Intelligent Network seeks to understand better these costs and benefits as they relate to Intelligent Network developments. The costs and benefits of unbundling ATM-based networks should likewise be examined. The analysis of ATM networks presented here suggests that unbundling is quite feasible, but a more detailed examination of the costs involved in providing various functions separately is warranted. The appropriate vehicle for examining these issues is the FCC's Notice of Inquiry process. The purpose of such an inquiry would be for the policy community to gain an appreciation of the possibilities for unbundling and the associated costs and benefits. Initiating an inquiry into this matter is appropriate even though ATM-based networks are still a few years away. At the very least, a public record of knowledge about ATM networks can be built, allowing interested parties access to information that is, for the most part, still confined to the technical journals. By opening a docket at this time, while the standards process is ongoing, public awareness of the issues at stake in standards development can be raised and can lead to greater, more representative participation in the process.

8.3.3 The Government's Role in Standards Development

After having determined the possibilities for open broadband networks and having weighed the associated costs and benefits, the FCC is then faced with the question of an appropriate course of action. Two issues are at stake: 1) does the government have a legitimate

---

[84] Mitchell Kapor, testimony at the FCC's en banc hearings, *Networks of the Future,* May 1, 1991. For an extended discussion of the merits of open networks, see Kapor and Berman (1992).

role in the development of telecommunications standards, and 2) how can policy considerations be embedded into telecommunications standards? The first issue revolves around the question of whether the setting of telecommunications standards can affect policy goals.[85] As Marcus (1991), in his analysis of technical standards and their policy implications, points out, some standards do and some do not have policy ramifications. Marcus cites the implementation of equal access for long-distance carriers as an example where technical decisions, made long before the divestiture of AT&T, complicated the realization of policy goals. Had an eventual opening of the long-distance market to competition been contemplated at the time that those decisions were made, the standards might have been different, and could have allowed for a smoother implementation. In the case of the FCC's policy of Open Network Architecture, the implications of B-ISDN standards are clear. The architecture of the networks—the technical decisions that will be made throughout standards process—will ultimately define what is meant by "Open Network Architecture." As Shefrin (1987) argues, the FCC, by leaving the details of ONA to be worked out in industry fora, effectively let the BOCs regulate themselves. If the FCC allows B-ISDN standards to develop without the benefit of clear goals for ONA, ONA will become only what the network standards allow it to be.

With regard to the second question, the method of injecting policy considerations into the standards process, one finds an interesting precedent. In 1983, the FCC opened a docket on Narrowband ISDN.[86] The Commission asserted an interest in the standardization of ISDN in part because of a concern that some of the ISDN concept seemed to involve classifications of services that would be difficult to reconcile with the *Computer II* distinction between basic and enhanced services.[87] The FCC's inquiry had an interesting result. At a meeting of the CCITT that took place after the FCC's *Notice of Inquiry*, the service classifications originally contemplated for

---

[85] For an overview of the policy significance of international standards setting processes, see U.S. Congress, Office of Technology Assessment (1992).

[86] In the Matter of Integrated Services Digital Networks, *Notice of Inquiry*, 94 FCC 2d 1289 (1983).

[87] The CCITT had originally discussed a distinction between "bearer" and "telecommunications" services that did not adequately approximate the basic/enhanced dichotomy of *Computer II*.

ISDN were modified in such a way as to approximate the basic/enhanced dichotomy of *Computer II*. The case provides an example of policy considerations entering into the standards deliberations and the Commission, satisfied with this result, concluded in its first report on the Inquiry, that the proceeding

> "... served the valuable purpose of focusing the attention of the industry and of government on the policy implications of ISDN planning."[88]

The Commission went on to reject the approach of subjecting ISDN planning to a formal rulemaking process and opted for a less direct form of intervention. The standards deliberations of voluntary organizations such as the T1 Committee of the Exchange Carrier Standards Association and of the State Department's advisory committee, the Commission concluded, offered ample opportunities for informal participation by the FCC and its staff. It is through these proceedings that the Commission chose to make known its policy goals and concerns.

To the extent that identified policy goals for ISDN were not compromised by the ISDN standards process, the informal, indirect approach of sensitizing the participants of the standards process to United States policy goals appears a reasonable solution to the problem of aligning standards to policy goals. It does not,'of course, ensure such an alignment. The process of standards setting at the CCITT is highly complex, and even arriving at a coherent U.S. position involves extensive negotiations. Should the FCC perceive that the indirect approach is not working effectively, it has the option of pressing its case directly to the State Department, who is ultimately responsible for the United States' negotiating position. It is important, in this discussion of government involvement in standards deliberations, that the involvement take the form of communicating policy goals rather than enforcing the adoption of particular technical standards. The B-ISDN standards process currently provides an opportunity for the development of open networks. As matter of policy, the Commission may wish to urge the CCITT, directly or indirectly, to choose architectural principles, as opposed to specific interface

---

[88] In the Matter of Integrated Services Digital Networks, *First Report*, 98 FCC 2d 289 (1985), at 289, para. 87.

characteristics, for B-ISDN that maximize flexibility and that maintain the potential for openness at many levels.

8.4 A NEW APPROACH TO UNIVERSAL SERVICE IS NEEDED

As the public telephone network undergoes a fundamental transformation into a public broadband communications network, the fate of the long-standing public policy of universal access to telephone service comes into question. As the network goes from a single technology, single service network to an all-purpose, multi-service network, will the principle of universal service be extended to cover integrated broadband networks? The universal service question can be divided into two separate issues. The first is the determination of which services are to be included in a basket of basic services that are to be made available to all people; the second is the means by which these services are to be made available. The second question, involving issues of subsidies, taxation, and regulatory incentives, can be analyzed without reference to the specific technology or services at issue and is thus not affected by the development of public ATM networks. It is in the area of defining universal service goals that the changes in the public network have profound impact.

Universal service was much easier to discuss in the days where the network essentially provided both one service (telephony), and one method of access (analog transmission over copper wires). To be sure, the definition of telephone service has undergone subtle changes, and access to single-party lines and touch-tone "dialing," to name two examples, are features or services that, in the past, were not part of universal service requirements (U.S. Department of Commerce, 1991). Despite the subtle changes in the definition of basic service, and the lack of a single, national statutory definition of basic service, universal service has, for the most part, been an issue of either having a connection to the telephone network or not having a connection to the telephone network.[89] As the range of services supported by the network and the number of methods of connection to the network grow, the consensus on what constitutes a basic service, one which is to be made available to all people, will surely break down.

The increasing amount of information available over the network and the increasing number of options for reaching that information lead to the concern that our society can become divided between information "haves" and "have-nots."[90] On the other hand, the proliferation of services and access methods allows for greater consumer choice. Lumping all communication services together into a basic package of services to be provided to all people ignores the real differences among communications users in their demand for certain communication services. Such an approach would have the undesirable effect of forcing consumers to pay for services that they do not want or cannot use. Any approach to defining basic services for universal provision must necessarily balance these countervailing concerns.

8.4.1 The Technology-Oriented Approach

One possible approach to defining universal service requirements is based on the technology of the network. As a network technology was developed and introduced successfully into the market, network providers would be required to pursue ubiquitous deployment. Such an approach would take the form, for example, of requiring that all people have access to

---

[89] Although the 1934 Communications Act does not discuss universal service and its meaning explicitly, the preamble of the act does contain a general policy goal, often interpreted as a mandate for universal service:

> "For the purpose of regulating interstate and foreign commerce in communication by wire and radio so as to make available, so far as possible, to all people of the United States a rapid, efficient, Nation-wide, and world-wide wire and radio communication service with adequate facilities at reasonable charges .... there is hereby created a commission to be known as the 'Federal Communications Commission,'" (47 U.S.C. 151).

[90] See statement of Congressman Edward J. Markey, Committee on Energy and Commerce, Subcommittee on Telecommunications and Finance, Hearings on Modified Final Judgment, 101st Cong. 1st Sess., (May 4, 1989), quoted in Kapor and Berman (1992).

Narrowband ISDN connectivity. As the network developed other capabilities that were deemed basic to communication, they too could be subject to universal service requirements.

One problem associated with any technologically-based approach is, of course, the tendency of technology to change rapidly. By the time that a new network technology has been implemented ubiquitously, it may be considered obsolete. A more fundamental shortcoming of the technology-oriented approach is the inadequacy of technology by itself. Network technology is a means to deliver services, and not a service in itself. For example, an N-ISDN connection to the network is not much use if there are few ISDN-based services to which one can connect. This scenario could arise if high-end users stimulate the provision of information services over broadband facilities, leaving fewer resources available to develop services for narrowband users. Furthermore, advances in the use of old technology can sometimes forestall the need for new technology in providing new services. While most discussions of residential video services focus on the need for fiber to the home or fiber to the curb, research on Asymmetrical Digital Subscriber Lines (ADSL) (Waring *et al*, 1991) and High-Speed Digital Subscriber Lines (HDSL) (Waring *et al*, 1991; Starr *et al*, 1991) has shown that standard quality television signals can be transmitted over copper subscriber loops.

While the focus may be better placed on access to services as opposed to access to network technologies, one should not underestimate the significance of the access bottleneck. The Intelligent Network principles, when more widely implemented, should minimize the problems associated with local provision of services. If services are bound to the software integrated into a local central office switch, then they can only be provided in areas served by capable central offices. If they are provided remotely, as is contemplated by the Intelligent Network concept, then the capabilities of the local central office are not critical to the offering of the services, so long as the local central office enables the user to reach the remote provider. The issue of access to services will focus on the access to the network—the on-ramp to the information age. The services to which a user will have access are limited by the nature of this on-ramp. A copper-based access, the ADSL/HDSL advances notwithstanding, will not provide access to some of the higher data rate services designed for delivery over optical fiber. This bottleneck issue complicates the problem of universal service because although a technology-oriented approach has considerable pitfalls, a service-based approach is ultimately limited by the technology employed in the subscriber access portion of the network.

8.4.2 The Content-Oriented Approach

An intriguing alternative to defining universal service requirements is a content-oriented approach (Hadden, 1991; Williams and Hadden, 1991). Certain types of information would be deemed socially beneficial (or necessary) and policies would be designed to ensure that all people could have access to this information. Hadden (1991) includes white pages directories, information on products and services (an "amplified Yellow Pages"), information about government activities and services, and educational services as candidates for universal access requirements.

This approach is fascinating in that it strikes at the heart of the purpose of information networks. How one gets to information is perhaps not as relevant, from a policy perspective, as whether one gets to it at all. This approach also marks the transition of the network from one that has primarily offered person-to-person communications to one that additionally connects people to information sources. In terms of universal access, providing enhanced forms of dial tone such as computer dial tone (direct digital connection to the network such as with ISDN) or the video dialtone proposed by the FCC, may not be sufficient to ensure that all people have access to the information that we, as a society, believe they should have.

Implicit in the content-oriented approach is, of course, the need to select the content that must be made available. This determination is a complex social issue and would likely invite political controversy. Even if a consensus can be achieved on the types of information to be made universally available, the question of appropriate information sources arises. Assuming that the government would not be the sole source of information, controversies may arise as to the objectivity of the information providers contracted for these mandated services. For example, who would be the providers of medical information? Local hospitals? Pharmaceutical manufacturers? Medical associations?

The content-oriented approach is further complicated by the fact that one cannot adequately separate form and content. Different media present information differently. A speech on the Senate floor, for example, could be stored as text and printed in the Congressional Record, could be broadcast over the radio, or could be made available on videotape. The text approach equates the "content" of the speech with the words, whereas the addition of audio and video components incorporates nuances such as intonation, gestures, and facial expressions that may alter one's interpretation of the words. Similarly, a health information service that offers high-resolution imagery, audio, and animated graphics cannot be equated to a service limited to unformatted, scrolling text characters. Content is inseparable from form and form is, to a large extent, limited by technology. A decision to make certain information universally available requires decisions about the form in which the information will be presented. User access to that information will depend on the form that is chosen and the ability of the user, given his or her technological capabilities, to process information in that form.

8.4.3 Summary

As the above analysis has shown, the problem of defining services that are to be universally available affords no simple answers. Any approach that focuses too closely upon network technology falls victim to the pace of technological change, yet no solution can be technology-free for the ability to reach information will always be a function of the network access capabilities available to the user. While the problem is challenging, it is also central to telecommunications policy and merits considerable attention. If we are to resolve the issue of information "haves" and "have-nots," we must work to develop a consensus as to the extent that the public network shall truly be public.

8.5 GENERAL CONCLUSIONS

An analysis of Asynchronous Transfer Mode technology shows that ATM-based networks can be regarded as service-independent communications platforms. With the introduction of ATM, the telephone network will be transformed to the point that telephone service will become just one of many communication services provided on top of a more fundamental network. Because telephone service and basic network provision will have different economic characteristics and because they may not share the same policy goals, telecommunications policy should reflect this conceptual separation of the telephone network from the more basic platform network. This separation of the two services need not be entirely conceptual—regulators may begin to implement this philosophy today by adopting an approach to network cost allocations that reflects this principle.

An examination of the structure of ATM-based broadband communications networks shows that their functions are both separable and modular. This functional architecture offers an opportunity for the furtherance of the Federal Communications Commission's policy of Open Network Architecture. The CCITT's Broadband ISDN standards project is ongoing. The degree to which the network will ultimately be open will depend strongly on the results of these standardization efforts. If the standardization process continues without regard to public policy goals, standards that preclude or complicate the implementation of open networks may result. It is therefore appropriate for the FCC to examine more closely the opportunities for open B-ISDN architecture and to participate in the B-ISDN standards process for the purpose of furthering its policy goals.

Finally, as the telephone network becomes an all-purpose broadband communications network, policy makers are challenged to develop new approaches to the policy of universal service. Any approach must recognize the real differences in demand for different services among different consumers without creating a society of information rich and information poor. This challenge is complicated by the ever changing technology of the network and by the close relationship between technology and the ability to gain access to information.

Glossary

AAL	ATM Adaptation Layer
ADSL	Asymmetrical Digital Subscriber Line
AIN	Advanced Intelligent Network
AM	Amplitude Modulation
ARPA	Advanced Research Projects Agency
ARPANET	Advanced Research Projects Agency Network
ASCII	American Standard Code for Information Interchange
AT&T	American Telephone and Telegraph Company
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband ISDN
BOC	Bell Operating Company
BRI	Basic Rate Interface
BSE	Basic Service Element
CAC	Connection Acceptance and Control
CCITT	International Telegraph and Telephone Consultative Committee
CD-ROM	Compact Disc Read Only Memory
CEI	Comparably Efficient Interconnection
CLP	Cell Loss Priority
CONAP	Coalition of Open Network Architecture Parties

CPE	Customer Premises Equipment
CPU	Central Processing Unit
DOS	Disk Operating System
ESP	Enhanced Service Provider
FCC	Federal Communications Commission
HDSL	High-Speed Digital Subscriber Line
IN	Intelligent Network
IP	Internet Protocol
ISDN	Integrated Services Digital Network
IWU	Inter-Working Unit
LAN	Local Area Network
LEC	Local Exchange Carrier
MCI	Microwave Communications, Inc.
MFJ	Modification of Final Judgment
MFS	Metropolitan Fiber Systems
MIT	Massachusetts Institute of Technology
MS-DOS	Microsoft Disk Operating System
MSO	Multiple Service Operator
N-ISDN	Narrowband ISDN
NREN	National Research and Education Network
NSFNET	National Science Foundation Network
ONA	Open Network Architecture
OSI	Open Systems Interconnection
PAD	Packet Assembler/Disassembler
PCA	Protective Coupling Arrangement
PCM	Pulse Code Modulation

PCN	Personal Communications Network
PCS	Personal Communications Services
POP	Point-of-Presence
POTS	Plain Old Telephone Service
PRI	Primary Rate Interface
PSTN	Public Switched Telephone Network
QOS	Quality of Service
SAGE	Semiautomatic Ground Environment
SAR	Segmentation and Reassembly
SCP	Service Control Point
SEAL	Simple and Efficient Adaptation Layer
SMDS	Switched Multimegabit Data Service
SONET	Synchronous Optical Network
SS7	Signaling System 7
TCI	Tele-Communications, Inc.
TMN	Telecommunications Management Network
VCI	Virtual Channel Identifier
VPI	Virtual Path Identifier

Bibliography

Aaron, M. R. and Maurizio Dècina (1991). Asynchronous Transfer Mode or Synchronous Transfer Mode or Both? *IEEE Communications Magazine*, 29(1): 10-13.

Ali, M. Irfan (1992). Frame Relay in Public Networks. *IEEE Communications Magazine*, 30(3): 72-78.

Anania, Loretta and Richard Jay Solomon (October 1988). Flat—The Minimalist B-ISDN Rate. Paper presented at Sixteenth Annual Telecommunications Policy Research Conference.

Appleton, James (1991). Performance Related Issues Concerning the Contract Between Network and Customer in ATM Networks. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends*. North-Holland, Amsterdam.

Aprille, Thomas J. (1990). Introducing SONET into the Local Exchange Carrier Network. *IEEE Communications Magazine*, 28(8): 34-38.

Baldwin, Bill (1991). Integrating ISDN Lines for Financial Users. *Telecommunications*, 25(6): 34-39.

Ballart, Ralph and Yau-Chau Ching (1989). SONET: Now It's the Standard Optical Network. *IEEE Communications Magazine*, 27(3): 8-15.

Beau, O., J. M. Silva, and H. Verhille (1990). Network Aspects of Broadband ISDN. *Electrical Communication*, 64(2/3): 139-146.

Bellcore Special Report SR-NWT-001937 (February 1991). *National ISDN-1*. Bell Communications Research, Morristown, NJ.

Bellcore Special Report SR-TSY-001681 (June 1990). *Bellcore Fiber In The Loop (FITL) Architecture Summary Report*. Bell Communications Research, Morristown, NJ.

Berman, Roger K. and John H. Brewster (1992). Perspectives on the AIN Architecture. *IEEE Communications Magazine*, 30(2): 27-32.

Bolter, Walter G. and James W. McConnaughey (1991). Innovation and New Services. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

Bonatti, M., F. Casali, and N. Heenan (1991). Evolution to B-ISDN: Overview and Preliminary Guide-lines. In Bonatti, M., F. Casali, and G. Popple, editors, *Integrated Broadband Communications, Views from RACE*. North-Holland, Amsterdam.

Bornholz, Robert and David S. Evans (1983). The Early History of Competition in the Telephone Industry. In Evans, David S., editor, *Breaking Up Bell: Essays on Industrial Organization and Regulation*. North Holland, New York.

Botein, Michael (1991). Regulatory Status: A Preliminary Inquiry. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues*. North-Holland, New York.

Bryan, J. Shelby (1991). PCN: Prospects in the United States. *Telecommunications*, 25(1): 54-56.

Burkhart, Lori A. (1992). Local Telephone Monopoly—The Beginning of the End. *Public Utilities Fortnightly*, 129(3): 32-34.

Buzacott, Alan (1990). *The Development of Broadband Telecommunications Standards*. Unpublished Master's Thesis, Massachusetts Institute of Technology.

Cavanagh, James P. (1992). Applying the Frame Relay Interface to Private Networks. *IEEE Communications Magazine*, 30(3): 48-64.

CCITT Recommendation I.121 (1991). *Broadband Aspects of ISDN*. Geneva.

CCITT Recommendation I.150 (1991). *B-ISDN Asynchronous Transfer Mode Functional Characteristics*. Geneva.

CCITT Recommendation I.211 (1991). *B-ISDN Service Aspects*. Geneva.

CCITT Recommendation I.311 (1991). *B-ISDN General Network Aspects*. Geneva.

CCITT Recommendation I.327 (1991). *B-ISDN Functional Architecture*. Geneva.

CCITT Recommendation I.413 (1991). *B-ISDN User-Network Interface*. Geneva.

CCITT Recommendation I.610 (1991). *OAM Principles of the B-ISDN Access*. Geneva.

CCITT Recommendation M.30 (1989). *Principles for a Telecommunications Management Network*. CCITT Blue Book, Volume IV, Fascicle VI.1, Geneva.

Cerf, Vinton G. (1991). Networks. *Scientific American*, 265(3): 72-81.

Cheung, Nim K. (1992). The Infrastructure for Gigabit Computer Networks. *IEEE Communications Magazine*, 30(4): 60-68.

Chiddix, James A. (1990). Fiber Backbone Trunking in Cable Television Networks: An Evolutionary Adoption of New Technology. *IEEE LCS*, 1(1): 32-27.

Choi, Don (1990). Multi-Level Precedence in B-ISDN. *Proceedings—IEEE Military Communications Conference*. Monterey, CA.

Codding, George A. and Anthony M. Rutkowski (1982). *The International Telecommunication Union in a Changing World.* Artech House, Dedham, MA.

Coon, Horace (1939). *American Tel & Tel: The Story of a Great Monopoly.* Longmans, Green and Co., New York.

De Prycker, M., J. L. Paul, and A. Campos (1990). Broadband National Experiments in Belgium, France, and Spain. *Electrical Communication,* 64(2/3): 218-228.

De Prycker, Martin (1991). *Asynchronous Transfer Mode: Solution for Broadband ISDN.* Ellis Horwood, London.

(1992). ATM Switching on Demand. *IEEE Network,* 6(2): 25-28.

Debuysscher, P., J-P. Glon, and M. Langenbach-Belz (1990). Evolution Towards Broadband. *Electrical Communication,* 64(2/3): 260-268.

Dobrowski, G. H., G. H. Estes, D. R. Spears, S. M. Walters (1990). Implications of B-ISDN Services on Network Architecture and Switching. *Proceedings of the Thirteenth International Switching Symposium,* Stockholm.

Dobrowski, G. H., M. Kerner, D. R. Spears, D. S. Wilson (1990). Evolving the Network Toward B-ISDN. *Proceedings of the Thirteenth International Switching Symposium,* Stockholm.

Dowling, Michael (1991). Value-Added Services: Regulation and Reality in the USA. *Telecommunications Policy,* 15(6): 509-518.

Egan, Bruce L. (1991). Conflict Between Public Policy and Trends in Technology. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues.* North-Holland, New York.

Electronic Industries Association (1991). *Electronic Market Data Book.* Washington, DC.

(1992). *Electronic Market Data Book.* Washington, DC.

Engineer, Carl P. (1991). Can PON Go Broadband? *Telephony,* 220(2): 30-35.

Fink, Robert L. and Floyd E. Ross (1992). Following the Fiber Distributed Data Interface. *IEEE Network,* 6(2): 50-55.

Frame, Mike (1990). Broadband Service Needs. *IEEE Communications Magazine,* 28(4): 59-62.

Geeslin, Bailey M. (1991). Innovation and New Services. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era.* Columbia University Press, New York.

Giancarlo, Charles H. (1992). ATM Forum Update: Accelerating Interoperability. *Telecommunications,* 26(7): 48.

Goodman, David J. (1991). Trends in Cellular and Cordless Communications. *IEEE Communications Magazine,* 29(6): 31-40.

Greenwald, Bruce C. N. (1991). Advances in Network Technology. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

Hac, Anna and Hasan B. Mutlu (1989). Synchronous Optical Network and Broadband ISDN Protocols. *Computer*, 22(11): 26-34.

Hack, David and Stephen J. Downs (1991). Should the "Baby Bells" Be Allowed to Manufacture? *CRS Issue Brief*. Congressional Research Service, Washington, DC. Updated Regularly.

Hadden, Susan G. (March 1991). Regulating Content as Universal Service. Working Paper, Policy Research Project: "Universal Service for the Twenty-First Century," The University of Texas at Austin.

Hellemans, Patrick, Hubert Decuypere, and Rik Verstraete (1990). A New Approach to Service Intelligence Supported by ATM. *IEEE Global Telecommunications Conference and Exhibition*, vol. 1. IEEE, Piscataway, NJ.

Horowitz, Edward, Robert Hubbard, Keith Pennington, and Andrew Lippman (1992). Digital Television in the Home and Workplace. Presented at the *MIT Communications Forum*, March 12.

Huber, Peter (1987). *The Geodesic Network: 1987 Report on Competition in the Telephone Industry*. U.S. Government Printing Office, Washington, DC.

Jain, Bijendra N. and Ashok K. Agrawala (1990). *Open Systems Interconnection: Its Architecture and Protocols*. Elsevier, Amsterdam.

Kahin, Brian (1992). Overview: Understanding the NREN. In Kahin, B., editor, *Building Information Infrastructure*. McGraw-Hill, New York.

Kaiser, Peter (1991). Broadband Trends in the U.S. In Bonatti, M., F. Casali, and G. Popple, editors, *Integrated Broadband Communications, Views from RACE*. North-Holland, Amsterdam.

Kapor, Mitchell and Jerry Berman (1992). Building the Open Road: The NREN as Testbed for the National Public Network. In Kahin, B., editor, *Building Information Infrastructure*. McGraw-Hill, New York.

Kleinrock, Leonard (1992). The Latency/Bandwidth Tradeoff in Gigabit Networks. *IEEE Communications Magazine*, 30(4): 36-40.

Kung, H. T. (1992). Gigabit Local Area Networks: A Systems Perspective. *IEEE Communications Magazine*, 30(4): 79-89.

Li, Tingye (1983). Advances in Optical Fiber Communications: An Historical Perspective. *IEEE Journal on Selected Areas in Communications*, 3: 356-367.

Lobjinski, M., G. Horn, and M. Horn (1990). Realization of Broadband Data Services in ATM-Networks. *Electronic Circuits and Systems for Communication: International Zurich Seminar on Digital Communications*. IEEE, Piscataway, NJ.

Lucky, Robert W. (1992). The Changing World of Communications. Presented at the *MIT Communications Forum*, May 7.

Lyles, J. Bryan and Daniel C. Swinehart (1992). The Emerging Gigabit Environment and the Role of Local ATM. *IEEE Communications Magazine*, 30(4): 52-58.

Marcus, Michael J. (1991). Technical Standards and Their Policy Implications. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues*. North-Holland, New York.

Mathison, Stuart L. and Philip M. Walker (1970). *Computers and Telecommunications: Issues in Public Policy*. Prentice-Hall, Englewood Cliffs, NJ.

McGarty, Terence P. (1992). Alternative Networking Architectures: Pricing Policy, and Competition. In Kahin, B., editor, *Building Information Infrastructure*. McGraw-Hill, New York.

McGowan, William G. (1991). Policy Directions for the Future. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

McKnight, Lee (1987). The International Standardization of Telecommunications Services and Equipment. In Mestmäcker, E-J., editor, *The Law and Economics of Transborder Telecommunications*. Nomos Verlagsgesellschaft, Baden-Baden.

McQuilian, John (1991). Cell Relay: High-Speed Networks Take a Step Closer to the 21st Century. *Data Communications*, 20(11): 58-69.

Merritt, John (1992). The Future of Frame Relay. *TE&M*, 96(1): 33-35.

Meyerson, Michael I. (1991). Impending Legal Issues. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues*. North-Holland, New York.

Millman, S., editor (1984). *A History of Engineering and Science in the Bell System: Communications Sciences (1925-1980)*. AT&T Bell Laboratories, Murray Hill, NJ.

Minzer, Steven E. (1989). Broadband ISDN and Asynchronous Transfer Mode (ATM). *IEEE Communications Magazine*, 27(9): 17-24.

Mitchell, Bridger M. and Ingo Vogelsang (1991). *Telecommunications Pricing: Theory and Practice*. Cambridge University Press, Cambridge.

Nagelhout, Mary (1988). Open Network Architecture: State Commissions Comment on Equal Access Plans. *Public Utilities Fortnightly*, 122(2): 44-47.

(1990). Teleport Complaints Push for ONA in New York and California. *Public Utilities Fortnightly*, 126(4): 48-50.

(1991). More Competition on the "Information Age" Horizon. *Public Utilities Fortnightly*, 127(8): 47-49.

National Research Council (1989). *Growing Vulnerability of the Public Switched Networks: Implications for National Security Emergency Preparedness*. National Academy Press, Washington, DC.

Noam, Eli M. (1989). Network Pluralism and Regulatory Pluralism. In Newberg, P. R., editor, *New Directions in Telecommunications Policy*. Duke University Press, Durham, NC.

(1991). Network Integration Versus Network Segmentation. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues*. North-Holland, New York.

(1992). Beyond the Golden Age of the Public Network. In Sapolsky, H., R. J. Crane, W. R. Neuman, and E. M. Noam, editors, *The Telecommunications Revolution: Past, Present, and Future*. Routledge, London.

O'Neill, E. F., editor (1985). *A History of Engineering and Science in the Bell System: Transmission Technology (1925-1975)*. AT&T Bell Laboratories, Murray Hill, NJ.

Patel, Virat (1992). Broadband Convergence: A View of the Regulatory Barriers. *Telecommunications Policy*, 16: 98-104.

Pepper, Robert M. (1988). *Through the Looking Glass: Integrated Broadband Networks, Regulatory Policy and Institutional Change*. OPP Working Paper Series No. 24, Federal Communications Commission, Washington.

(1991). Players and Stakes. In Elton, M. C. J., editor, *Integrated Broadband Networks: The Public Policy Issues*. North-Holland, New York.

Phillips, Almarin (1991). Changing Markets and Institutional Inertia: A Review of US Telecommunications Policy. *Telecommunications Policy*, 15(1): 49-61.

Pool, Ithiel de Sola (1983). *Technologies of Freedom*. Harvard University Press, Cambridge, MA.

Radford, William H. (1962). MIT Lincoln Laboratory: Its Origin and First Decade. *Technology Review*. 64(3): 15-17, 40-42.

Ransom, M. Niel and Dan R. Spears (1992). Applications of Public Gigabit Networks. *IEEE Network*, 6(2): 30-40.

Rao, S. (1991). Access Architectures for Broadband ATM Networks in the Business Community. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends*. North-Holland, Amsterdam.

Reed, David P. (December 1991). Taking It All Apart: Principles of Network Modularity. Paper presented at the Conference on Private Networks and Public Objectives, Columbia Institute for Tele-Information.

Rey, R. F., editor (1984). *Engineering and Operations in the Bell System*. AT&T Bell Laboratories, Murray Hill, NJ.

Roberts, Lawrence G. (1978). The Evolution of Packet Switching. *Proceedings of the IEEE*, 66: 1307-1313.

Robinson, Glen O. (1991). Regulatory and Institutional Change. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

Robrock, Richard B. (1991). The Intelligent Network—Changing the Face of Telecommunications. *Proceedings of the IEEE*, 79: 7-20.

Rutkowski, Anthony M. (1985). *Integrated Services Digital Networks*. Artech House, Dedham, MA.

Sandbach, Jonathan (1991). Network Competition and Value-Added Services: What Should Follow ONP? *Telecommunications Policy*, 15(6):479-484.

Sandesara, Niranjan B., G. Ray Ritchie, and Barbara Engel-Smith (1990). Plans and Considerations for SONET Deployment. *IEEE Communications Magazine*, 28(8): 26-33.

Sato, Harumasa and Rodney Stevenson (1989). Telecommunications in Japan: After Privatization and Liberalization. *Columbia Journal of World Business*, 24(1): 31-41.

Sato, Ken-ichi, Hisaya Hadama, and Ikuo Tokizawa (1990). Network Reliability Enhancement With Virtual Path Strategy. *IEEE Global Telecommunications Conference and Exhibition*, vol. 1. IEEE, Piscataway, NJ.

Sato, Youichi and Ken-ichi Sato (1991). Virtual Path and Link Capacity Design for ATM Networks. *IEEE Journal on Selected Areas in Communications*, 9(1): 104-111.

Sawyer, W. D. M., J. H. Salladay, and M. W. Snider (1991). The Gateway Role in an Open-Access, High-Capacity Network. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends*. North-Holland, Amsterdam.

Schindler, G. E., editor (1982). *A History of Engineering and Science in the Bell System: Switching Technology (1925-1975)*. Bell Telephone Laboratories, Inc.

Schneider, Herbert (1990). The Concept of Virtual Paths and Virtual Channels in ATM-Networks. *Electronic Circuits and Systems for Communication: International Zurich Seminar on Digital Communications*. IEEE, Piscataway, NJ.

Shefrin, Ivan H. (1987). *Telecommunications Regulation in Enhanced Services: The Third Computer Inquiry*. Unpublished Master's Thesis, Massachusetts Institute of Technology.

Shumate, Paul W. (1989). Optical Fibers Reach into Homes. *IEEE Spectrum*, 26(2): 43-47.

Shumate, Paul W. and Richard K. Snelling (1991). Evolution of Fiber in the Residential Loop Plant. *IEEE Communications Magazine*, 29(3): 68-74.

Sikes, Alfred C. (1990). After Computer III: Picking Up the Pieces at the FCC. *Public Utilities Fortnightly*, 126(4): 31-33.

(1991). Policy Directions for the Future. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

Singer, Richard M. and David A. Irwin (1991). Personal Communications Services: The Next Technological Revolution. *IEEE Communications Magazine*, 29(2): 62-66.

Sirbu, Marvin A. (1991). Advances in Network Technology. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era*. Columbia University Press, New York.

(1992). The Struggle for Control Within the Telecommunications Networks. In Sapolsky, H., R. J. Crane, W. R. Neuman, and E. M. Noam, editors, *The Telecommunications Revolution: Past, Present, and Future*. Routledge, London.

Skoog, R. A. and A. R. Modarressi (1990). Alternative Issues for Network Signaling Transport in a Broadband Environment. *Computer Networks and ISDN Systems*, 20: 361-368.

Smith, D.G. and D.C. Pitt (1991). Open Network Architecture: Journey to an Unknown Destination? *Telecommunications Policy*, 15: 379-394.

Solomon, Richard Jay (February 1989). Past & Future Perspectives on Communications Infrastructure. Presented at "Integrated Broadband Networks 2," Columbia University.

Stallings, William (1987). *Local Networks: An Introduction (2nd Edition)*. MacMillan, New York.

(1990). CCITT Standards Foreshadow Broadband ISDN. *Telecommunications*, 24(3): 29-41.

Starr, Thomas J. J., David L. Waring, and Jean-Jacques Werner (1991). High Bit-rate Digital Subscriber Line (HDSL): An Expedient Broadband Access. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends*. North-Holland, Amsterdam.

Stavenow, B. and K. Sällberg (1991). A Flexible Structure of the SMDS User Access Network. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends*. North-Holland, Amsterdam.

Tanaka, Toshiki P., Eiichi Amada, and Yoshihiro Takiyasu (1989). A Consideration on ATM Technology in Private Networks. *IEEE Global Telecommunications Conference and Exhibition*, vol. 3: 1835-1840. IEEE, Piscataway, NJ.

Tanenbaum, A. S. (1981). *Computer Networks*. Prentice-Hall, Englewood Cliffs, NJ.

Taylor, Steven A. (1992). Frame Transport Systems. *IEEE Communications Magazine*, 30(3): 66-70.

Temin, Peter (1987). *The Fall of the Bell System: A Study in Prices and Politics*. Cambridge University Press, Cambridge.

(1992). Did Regulation Keep Pace with Technology? In Sapolsky, H., R. J. Crane, W. R. Neuman, and E. M. Noam, editors, *The Telecommunications Revolution: Past, Present, and Future*. Routledge, London.

Teske, Paul and John Gebosky (1991). Local Telecommunications Competitors: Strategy and Policy. *Telecommunications Policy*, 15: 429-436.

U.S. Congress, Office of Technology Assessment (1990). *Critical Connections: Communication for the Future*. U.S. Government Printing Office, Washington, DC.

(1991). *The 1992 World Administrative Radio Conference: Issues for U.S. International Spectrum Policy—Background Paper.* U.S. Government Printing Office, Washington, DC.

(1992). *Global Standards: Building Blocks for the Future.* U.S. Government Printing Office, Washington, DC.

U.S. Department of Commerce, National Telecommunications and Information Administration (1991). *The NTIA Infrastructure Report: Telecommunications in the Age of Information.* NTIA Special Publication 91-26.

Veeraraghavan, M., D. M. Rouse, and R. Kapoor (1991). Signaling Architectures and Protocols for Broadband ISDN Services. In Lemstra, W., editor, *Telecommunication Access Networks: Technology and Service Trends.* North-Holland, Amsterdam.

Verbiest, Willem, Luc Pinnoo, and Bart Voeten (1988). The Impact of the ATM Concept on Video Coding. *IEEE Journal on Selected Areas in Communications,* 6: 1623-1632.

Walters, Stephen M. (1991). A New Direction for Broadband ISDN. *IEEE Communications Magazine,* 29(9): 39-42.

Waring, David L., Joseph W. Lechleider, and To Russell Hsing (1991). Digital Subscriber Line Technology Facilitates a Graceful Transition from Copper to Fiber. *IEEE Communications Magazine,* 29(3): 96-104.

Wiley, Richard E. (1991). Regulatory and Institutional Change. In Cole, B. G., editor, *After the Breakup: Assessing the New Post-AT&T Divestiture Era.* Columbia University Press, New York.

Williams, Frederick and Susan Hadden (1991). On the Prospects for Redefining Universal Service: From Connectivity to Content. Working Paper, Policy Research Project: "Universal Service for the Twenty-First Century," The University of Texas at Austin.

Zerbiec, Timothy G. (1992). Considering the Past and Anticipating the Future for Private Data Networks. *IEEE Communications Magazine,* 30(3): 36-46.